Identification and characterization of phage displayed, SC binding peptides, with potential for use in pIgR mediated mucosal targeting

Anders Sandvik

Thesis for the degree of Cand.scient in Molecular Cell Biology

Division of Cell and Molecular Biology,
Department of Molecular Bioscience,
University of Oslo
2003

CONTENTS

Preface

This thesis attempts to communicate basic knowledge about the immune system and how it can be modulated for potential use in modern medicine. Furthermore, based on that introduction, the thesis presents recent research related to mucosal immunology. The thesis comprises four separate, though interrelated parts:

- *General outline of the Immune system*
- *Antibody engineering*
- *Introduction*
- *Manuscript*

Part I, *General outline of the Immune system*, gives a brief introduction to the components, the organization, and effector mechanisms of the immune system. This part is intended for readers without training in immunology. My aim is to provide the reader with basic insight in the importance of the immune system and the prospective of capitalizing on our native defence systems in modern medicine. An introduction to immunological terminology and basic comprehension of the immune-biology is clearly essential for efficient communication of the ideas, results and discussion presented in this thesis.

Part II, *Antibody engineering*, gives an introduction to promising technology based on engineering and alternative production of antibodies. Recombinant antibodies and fragments of such show immense potential in therapeutic and diagnostic applications and as research tools. The available technology, prospects for the future and the limitations of this technology is given some attention as the research group, from which this work originates, has its focus on antibody modulation. Thus, a brief introduction to this exciting topic facilitates the communication and understanding of the work presented herein.

Part III, *Introduction*, gives a focused introduction to mucosal immunity, the use of phage display libraries in general and fUSE5 based systems in particular. Furthermore targeting of mucosal sites via the polymeric Ig-receptor is discussed. The theoretical basis for the work presented in the enclosed manuscript is dealt with in some detail in this section.

These three sections are based largely on recent review articles. Interested readers are referred to these for further background information and links to original work.

Part IV, Manuscript, concludes this thesis. ”*Identification and characterization of phage displayed, SC binding peptides, with potential for use in pIgR mediated mucosal targeting*”, presents the work I have conducted for the academic degree of Cand.scient.

Acknowledgement

The work presented in this thesis was carried out at the Division of Cell and Molecular Biology, Department of Molecular Bioscience, University of Oslo, from September 2001 to December 2003 for the degree of Cand.scient.

I would like to thank Professor Inger Sandlie for giving me the opportunity to become an immunologist. Inger has provided interesting and challenging tasks accompanied by expert guidance that has allowed me to grow both as a scientist and as a person during my stay in her research group. Furthermore I would like to thank her for inviting colleagues and students onto the social venue; be it dinners at her house or week-end escapes to the "residence" at Blefjell.

I also would like to thank Dr. Vigdis Lauvrak for invaluable laboratory guidance, technical expertise and fruitful discussions. Vigdis has been my day-to-day supervisor and project manager. Thank you for always being available and willing to help.

Thanks to all the members of the Sandlie lab-group for crating a stimulating, motivating and friendly environment. Especially, I would like to thank Gøril Berntzen for her effort in assisting me in my lab work. Also a special thanks to Elin T. Johansen, for expert insight in common frustrations and for rewarding discussions, both scientific and non-scientific.

Thanks to Thomas L. Kongsvik, for assistance with fluorescence microscopy.
Also thanks to Dr. Finn-Eirik Johansen and Ranveig Braathen, for their help on fluorescence measurements, inhibition- and transcytosis studies, for reagents and for motivating discussions.

Thanks to all my friends and fellow students at "study-room" 3525 for making the last two years of my life, as a student, a memorable period of a golden era.

Last but not least, I would like to thank my parents Kari and Odd Helge for supporting me, for believing in me and for having patients with me.

Oslo, December 2003

Anders Sandvik

Abbreviations

Aa, Amino acid

Ab(s), Antibody (ies)

ADCC, Antibody-dependent cell-mediated cytotoxicity

APC(s), Antigen presenting cell(s)

BALT, Bronchial-associated lymphoid tissue

BCR(s), B-lymphocyte receptor(s)

BPI, Bacterial permeability inducing protein

C, Constant region

CD, Clusters of differentiation

CDR(s), Complementarity-determining region(s)

C_H, Heavy chain constant region

C_L, Light chain constant region

CR(s), Complement receptor(s)

C-terminal, COOH-terminal

CTL(s), Cytotoxic effector T-lymphocyte(s)

D, Diversity region

DC(s), Dendritic cell(s)

D_H, Heavy chain diversity region

Fab, Fragment antigen binding

FcγRs, IgG Fc receptors

Fc, Fragment crystallisable

FcRn, Neonatal Fc receptor

FcRs, Fc receptors

Ff, F-pilus dependent filamentous phages

fSC, Free SC

Fv, Fragment variable

GALT, Gut-associated lymphoid tissue

HLA, Human leukocyte-associated antigens

Ig(s), Immunoglobulin(s)

IL(s), Interleukin(s)

ITAM(s), Immunoreceptor tyrosine-based activation motif(s)

J, Joining segment

J_H, Heavy chain joining segment

J_L, Light chain joining segment

mAbs, Monoclonal Abs

MAC, Membrane attack complex

MALT, Mucosa associated lymphoid tissue

MASP(s), MBL-associated serine protease(s)

MBL, Mannan-binding lectin

MC(s), Mast cell(s)

MDCK, Madin-Darby canine kidney

MHC, Major histocompatibility complex

NK cell(s), Natural Killer cell(s)

N-terminal, NH2-terminal

PALS, Periarteriolar lymphoid sheath

PAMP(s), Pathogen associated molecular pattern(s)

PCR, Polymerase chain reaction

pIgA, Polymeric IgA

pIgR, Polymeric immunoglobulin receptor

pIgs, Polymeric Igs

RF, Replication form

sAbs, Secretory Abs

SC, Secretory component

scFv, Single chain Fv

sIgA, Secretory IgA

sIgM, Secretory IgM

ssDNA, Single stranded DNA

TBR, Target binding region

TCR(s), T-lymphocyte receptor(s)

TLR(s), Toll-like receptors

V, Variable region

V_H, Heavy chain variable region

V_L, Light chain variable region

General outline of the Immune system

Our body is facing a constant challenge against disease causing agents, collectively known as pathogens. Four broad categories of pathogens are now recognized. These are viruses, bacteria, pathogenic fungi, and other relatively large and complex eukaryotic organisms collectively termed parasites. Every second we exist; some opportunistic invader (bacteria, virus, fungi or parasites) is trying to make a meal or breading ground of our body. If properly functioning, our immune system keeps these relentless invaders at bay with its incredible array of weapons. Distributed throughout the body, the immune system represents our "armed forces."

The immune system is a complex network of lymphoid tissue, cells, humoral factors and cytokines that interact to eliminate foreign invaders, pathogens and altered self, i.e. cancer. Host defence against the broad variety pathogens constantly present in our environment, requires dramatically different responses, depending on the character of the pathogen and on the tissue under attack. Central to the immune system's ability to mobilize a response to an invading pathogen is its sophisticated biologic ability to distinguish self from non-self and altered-self. We have evolved both innate and adaptive mechanisms to respond to and eliminate pathogens. Both of these branches of the immune system, further discussed in the next sections, rely on self-, non-self discrimination. A key feature of the adaptive response is that it produces long lived cells that can, upon re-challenge with the antigen, re-express effector mechanisms. This immunological memory provides the host with a specific and efficient defence when it encounters the pathogen a second time. It is of critical importance, that the powerful, destructive effector mechanisms of the immune system are not attacking the host's own cells. Avoidance of destruction of self-tissues is referred to as self tolerance. The broad class of autoimmune diseases are caused by failure of self-tolerance.

Innate immunity

The term innate immunity encompasses granulocytes, monocytes, macrophages, dendritic cells, mast cell, natural killer cells, complement, cytokines and acute phase proteins. Physical barriers such as epithelial cell layers and mucosal surfaces in the gut, respiratory- and urogenital tract are also sometimes considered part of the innate immune system.

Chemical- and antimicrobial barriers are also included in some definitions of innate immunity. These highly conserved defensive measurements of the immune system have evolved in even the simplest animals.

Cells and cellular effector mechanisms of the innate immune system

The innate immune response lacks specificity, it is rapid and is considered to be the first line of defence. Immune responses depend largely upon the activities of white blood cells, or leukocytes. All leukocytes derive ultimately from the same hematopoietic stem cells. The different leucocytes are classified both by morphology and by surface phenotype as determined by registered differentiation antigens known as clusters of differentiation (CD). Innate immunity largely involves cells derived from myeloid progenitors.

Granulocytes, also called polymorphonuclear leukocytes, are a diverse collection of white blood cells including the phagocytic neutrophils. **Neutrophils**, short-lived, migrating and well-armoured cells are possibly the most important cellular, component of the innate immune system. They detect the presence of pathogens through germ-line encoded receptors that recognize pathogen associated molecular patterns (PAMPs). In vertebrates, the best characterized class of these receptors are Toll-like Receptors (TLRs) (reviewed in Takeda et al., 2003; Underhill, 2003). Interaction between PAMPs and the surface receptors of the phagocytic cells triggers the cells to engulf the pathogen they have encountered. Phagocytosis by neutrophils involves additional important receptor classes: Fcγ receptors (FcγRIIA /CD32, and FcγRIIIB/CD16) and complement receptors (CR1 /CD35 and CR3/CD11b/CD18 integrin). Neutrophils hold an impressive arsenal of antibiotic enzymes, such as acidic hydrolases, myelo peroxidase and lysozyme and antibacterial proteins such as defensin and bacterial permeability inducing (BPI) protein. They also possess membrane bound compartments containing highly reactive, cytotoxic oxygen species, such as hydrogen peroxide and singlet oxygen used to fight invaders. Neutrophils are both a target and a source of various cytokines, chemokines, and growth factors (Witko-Sarsat et al., 2000, review). Cytokines are secreted proteins that affect the behaviour of other cells that bears receptors for them. Cytokines produced by white blood cells mainly work on other leucocytes and are collectively called interleukins (ILs). Chemokines are chemo attractants that recruit cells with chemokine receptors.

Circulating **monocytes** differentiate to become macrophages when they leave the blood and migrate towards sites of infection or inflammation. **Macrophages** are widely distributed, mononuclear, phagocytic cells. They are equipped with FcγReceptors, complement receptors and TLRs that enable them to recognize and engulf pathogens. Macrophages show microbicidal activity much like neutrophils. In addition they produce cytotoxic amounts of nitric oxide (Panaro et al., 2003). Early after an infection or tissue damage, activated macrophages release cytokines and chemokines, giving them an immuno-modulating and regulatory role in adaptive immune responses.

Dendritic cells (DCs) are a third phagocytic cell type. They are bone marrow-derived migratory cells that are present in virtually all organs of the body. Their major role is to sample micro environmental antigens, and then to migrate to regional lymph nodes where they present processed antigenic peptides to T-cells and initiate adaptive immunity. DCs produce, release and respond to cytokines. Microbial antigens can stimulate myeloid DCs to induce either Th1 or Th2 immune responses (reviewed in Upham, 2003). The phagocytes of the innate immune system provide a first line of defence against many common micro-organisms, they are essential for the control of common bacterial infections and go on to initiate and orchestrate later adaptive responses. Phagocytosis and killing of pathogens is greatly enhanced if the particle to be eliminated is first opsonised. That is if the pathogen is marked for destruction by bound antibodies (Abs) or surface bound soluble proteins such as the ones comprising the complement system. Phagocytosis of opsonised pathogens is mediated by Fc-receptors or complement receptors on the surface of the phagocytic cells.

Mast cells (MCs), most widely studied in the context of allergic disease and hypersensitivity reactions, differentiate in tissue and are mainly located near small blood vessels. When activated, they release substances that affect vascular permeability and induce inflammation (e.g. histamine, serotonine, leucotrienes and prostaglandins). MCs play a critical role in host defence against bacterial infection. They recognize the products of bacterial infection through several surface receptors. In addition to its protective role, mast cells have homeostatic functions. The cytokines and chemokines produced by mast cells in response to pathogens are known to profoundly alter the nature of the innate immune response and its effectiveness in eliminating infection facilitate subsequent tissue repair (reviewed in Boyce, 2003; Marshall et al., 2003).

Basophils, found in circulation, are morphologically and functionally similar to mast cells. Like mast cells they posses high affinity receptors for IgE ($Fc_{\varepsilon}RI$). Cross-linking of the receptors leads to degranulation and release of preformed effector molecules.

Eosinophils mediate host protection against larger parasites, such as nematodes, that can not be phagocytosed. Eosinophils express low affinity receptors ($Fc_{\varepsilon}RII$) that interact with pathogens opsonised with IgE. Binding to IgE leads to degranulation and release of cytotoxic substances onto the pathogen surface.

Natural Killer cells (NK cells) a lineage of lymphoid cells that lack antigen specific receptors are part of the innate immune system. These cells circulate in the blood as large lymphocytes with distinctive cytotoxic granules. NK cells scan the host's cells and extracellular environment for abnormalities by means of both inhibiting- and activating surface receptors. Their cytotoxic activity is inhibited by binding to self- MHC I molecules on the host's cell surface. Down regulated or missing MHC I on the host's cells, i.e. caused by viral infection, trigger the NK cells to kill the target by inducing apoptosis in the infected self-cell. This selection mechanism known as "missing self" activation, allows for NK cells to recognize and kill abnormal cells, for example some tumour cells and virus-infected cells (Karre, 1981; Karre, 2002, review). NK cells also hold immunoglobulin receptors (FcRs). Binding of Ab coated antigens mediates NK-cells to induce antibody-dependent cellular cytotoxicity. NK cells are thought to be important in the innate immune defence against intracellular pathogens and altered self (reviewed in Middleton et al., 2002).

The complement system: an overview

Complement was initially discovered as heat-sensitive components of normal plasma that enhanced antibody killing of some bacteria. This activity was said to 'complement' the antibacterial activity of antibodies, hence the name. Although first discovered as an arm of adaptive immune responses, complement can be activated early in infection independent of the presence of antibodies. It is now evident that complement first evolved as part of the innate immune system, where it still plays an important role (Nonaka, 2001, review). The complement system (see figure 1) is composed of more than 30 plasma and cell surface proteins. The main goals of complement activation is; firstly to mark targets permanently for destruction. Secondly, to recruit other proteins and cells that facilitates target destruction.

Thirdly, to participate directly in destroying pathogens by formation of the membrane attack complexes (MAC) on the target, thereby inducing osmotic lysis. The initiation of the complement system triggers a cascade of reactions. The early components of the cascade are serine protease pro-enzymes, also known as zymogens that are sequentially activated by limited proteolytic cleavage, thereby amplifying the initial signal. The cascade can be initiated via three known pathways. These pathways depend on different molecules for activation, but they converge to generate the same set of effector molecules.

Immune-complexes formed on the pathogen surface by binding of antibodies (IgM and IgG) to antigen surfaces initiates **the classical pathway**. Binding of multiple Fc-regions to the globular heads of C1q, a subcomponent of C1, a component of the complement system, initiates a cascade of reactions. Sequential activation of complement components C1, C4 and C2 leads to the generation of the key enzyme C3 convertase. C3 convertase, the enzyme that cleaves C3 to C3a and C3b, is the common denominator of all three pathways that subsequently goes on to activate the later steps of the cascade. Host cells express protein regulators of complement activation (RCA), which inhibit the complement cascade at the C3/C5 convertase step. On foreign surfaces on the other hand, the C3 convertase is free to catalyse substrate cleavage. C3b bound to the C3 convertase forms a C5 convertase. C5b fragments bind C6 and C7, to form hydrophobic C5b67 complexes that attract C8 molecules and insert into the cell membrane. Finally, further recruitment of C9 molecules leads to membrane lesion and cell lysis by formation of a self assembling pore-forming complex known as the membrane attack complex (MAC). During complement activation, fragments that act as opsonins and chemo attractants are generated. C4b and C3b, deposited on the cascade triggering agent, function as opsonins by binding the complement receptor 1 (CR1) on phagocytic neutrophils, monocytes and macrophages, leading to enhanced phagocytosis and clearance of pathogens. Inflammatory molecules, such as C3a, C4a and C5a induce smooth muscle contraction and increased vascular permeability, which recruit Abs and complement molecules to the infection site. In addition, C5a functions as a chemoattractant that directs the migration of phagocytotic cells towards inflammatory tissue.

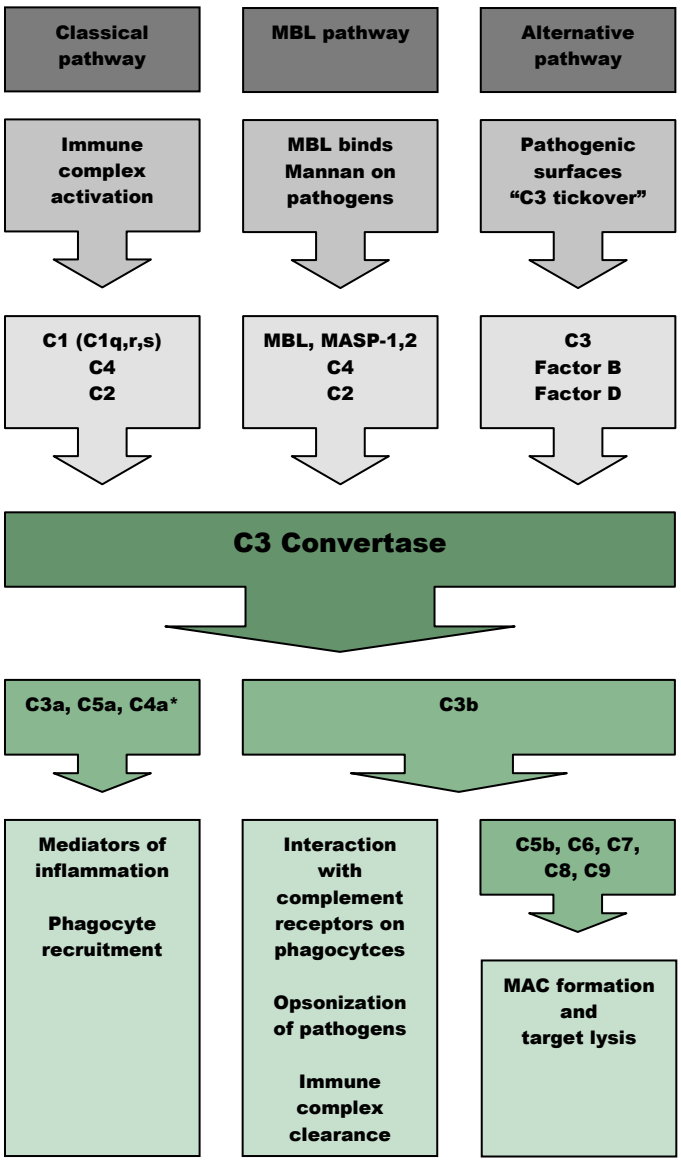


Figure 1. Overview of the main components and effector functions of the complement system.

The classical pathway is initiated by the formation of immune complexes on the pathogens surface. Multimeric binding of antibody Fc regions to the globular heads of subcomponent C1q activates proteolysis of C1. C1 subsequently cleaves C4 and C2 to form the classical C3 convertase.

The mannose binding lectin pathway (MBL) is activated through contact between mannan-containing microbes and MBL. This interaction activates mannan-binding lectin-associated serine protease 1 and 2 (MASP-1, 2) leading to C2 and C4 cleavage. Once again, C3 convertase is formed.

The alternative pathway is initiated by interaction between complement inhibitors and microbial antigens that cancel out the inhibitory signals. This allows for auto activation of the cascade at the C3 level to occur. C3 interacts with factor B and factor D to generate the alternative pathway C3 convertase.

The early events of all three pathways involve a series of cleavage reactions that culminate in the formation of a C3 convertase. Generation of C3 convertase is the point at which the three pathways converge and the main effector functions of the complement system are generated.

All three C3 convertases cleave component C3 to form C3a and the larger C3b. C3a increase vascular permeability allowing proteins, such as Abs, and migrating phagocytes to penetrate the tissue. It also works as a chemo attractant. C3b is deposited on the activating pathogenic surface or immune complex. This opsonizes the activating antigen for enhanced phagocytosis and initiates the formation of a membrane-attack complex (MAC).

C5a and C5b are generated by cleavage of C5 by a C5 convertase formed by C3b bound to the C3 convertase (not shown in this simplified diagram). C5a, like C3a, is a powerful peptide mediator of inflammation.

Formation of C5b triggers the late events in which the terminal components of complement, C6, C7, C8 and C9 spontaneously assemble into a membrane-attack complex. MAC forms a pore structure that damages the membrane of certain pathogens and there by kill them by osmotic lysis.

C4a, generated by the cleavage of C4, prior to the formation of C3 convertase, during the early events of the classical pathway, hence the *; it is a relatively weak peptide mediator of inflammation. Similarly, C4b, the large cleavage fragment of C4 (not shown in this simplified diagram), is a weak opsonin.

The diagram is based on figure 2.8 in Immunobiology, 5th ed, Janeway, et al.

The alternative pathway is independent of Abs and is activated by microbial surface structures that counterbalance inhibitors of spontaneous complement activation. The alternative pathway is initiated by spontaneous hydrolysis of the internal thioester bond within plasma C3, also known as the C3 'tick over' mechanism. The larger cleavage product C3i, a structural and functional analogue of C3b, associates with a protein of the alternative pathway called factor B. Factor D, another protein of the alternative pathway, cleaves the C3b-factor B complex to generate the alternative pathway C3 convertase, that subsequently goes on to activate later steps in the cascade. The alternative pathway also amplifies the classical pathway of complement activation. The spontaneously activated complement products indiscriminately bind both host and foreign surfaces and may form lytic membrane attack complexes. As mentioned shortly above, host cells are protected from complement lysis by membrane bound and soluble proteins that inhibit MAC formation, whereas pathogens lack this protection.

The so called **lectin pathway** is triggered by microbial cell wall components containing mannans and other related carbohydrates. The interaction between pathogenic mannan and serum mannan-binding lectin (MBL), a molecule that resembles C1q, activates pro-enzymes known as MBL-associated serine protease 1 and 2 (MASP-1, MASP-2). MASP-1and 2 form a protease analogous to the activated C1 protease found in the classical pathway that goes on to activate the rest of the complement cascade (reviewed in Walport, 2001a; Walport, 2001b).

Phagocytic cells also play an important role in adaptive immune response by taking up pathogens, processing them with proteolytic enzymes and eventually displaying pathogenic peptide fragments that can activate T-cell responses. Additional related cells, including Langerhans cells of epidermis, Kupfer cells in liver, Microglial cells in the central nervous system and a broad class of DC's, do all express MHC I and MHC II surface molecules that permit them to display internalized and processed pathogenic peptide-digests to T-cell receptors (TCRs) on T-lymphocytes (T-cells are discussed in further detail later on).
DC's appear to be the most effective antigen presenting cells (APCs). However, all MHC expressing cells are potential APCs if stimulated appropriately. The innate immune response makes a crucial contribution to the activation of adaptive immunity. The inflammatory response and release of molecular signals caused by complement activation, increases the flow of lymph containing antigen and antigen-bearing cells into secondary lymphoid tissue,

while complement fragments on microbial surfaces and induced changes in cells that have taken up pathogens provide signals that synergize in activating lymphocytes whose receptors bind to microbial antigens (reviewed in Cole and Morgan, 2003; Nielsen and Leslie, 2002). Antigen recognition, presentation and the details of effector mechanisms of adaptive immunity, is discussed in greater detail in the next section.

Adaptive immunity

Unlike the innate mechanisms, the adaptive reaction is highly specific for its target antigens. However, the adaptive mechanisms typically need several days before they become ready for combat. The adaptive responses are based on the antigen specific receptors expressed on the surface of T- and B-lymphocytes. The ability to recognize infinite numbers of antigens is enabled by somatic rearrangement of genes that encode T-cell receptors (TCRs) and immunoglobulins (Igs), as well as by the requirement for T cells to recognize antigens in the context of presentation by MHC molecules encoded within the major histocompatibility complex.

Lymphoid tissue; birth place and playground of lymphocytes

The lymphoid organs are organized tissues containing large numbers of lymphocytes in a framework of non-lymphoid cells. Lymphoid tissue can be divided into primary lymphoid organs, where lymphocytes are generated, and secondary lymphoid organs, where adaptive immune responses are initiated and where lymphocytes are maintained.

The primary lymphoid organs are the bone marrow and the thymus. The leucocytes of adaptive responses are derived from the same lymphoid progenitor cells in the bone marrow. Two major types of lymphocyte exists: B lymphocytes (B cells) and T lymphocytes (T cells or Thymocytes). B cells mature in the bone marrow, whereas T cells migrate to the thymus to mature. Upon maturation, the lymphocytes enter the bloodstream, from which they migrate to the secondary lymphoid organs.

Secondary lymphoid tissues include the lymph nodes, the spleen, and mucosa associated lymphoid tissue. The primary role of secondary lymphoid tissue is to facilitate interactions between antigens, antigen charged APCs or immune complexes and recycling lymphocytes.

Lymph nodes are highly organized structures located at the points of convergence of vessels of the lymphatic system, an extensive system of vessels that collects extracellular fluid from the tissues and returns it to the blood. This extracellular fluid, called lymph is produced continuously by filtration from the blood. It is transported in vessels known as lymphatic vessels. The lymphatic vessels collect fluid from the tissues and the lymph also carry antigen-bearing cells and antigens from infected tissues to the lymph nodes where they are trapped. B and T lymphocytes are segregated in the lymph nodes. B-cells are organized in follicles, surrounded by more diffusely distributed T-cells in what is referred to as T-cell zones.

The spleen collects antigen from the blood. Lymphocytes surround the arterioles entering the organ, forming areas known as white pulp. This inner region is divided into the periarteriolar lymphoid sheath (PALS), containing mainly T cells, and a flanking B-cell corona (reviewed in Chadburn, 2000).

Mucosa associated secondary lymphoid tissue (MALT) are frequently subdivided in to gut-associated lymphoid tissues (GALT) and bronchial-associated lymphoid tissue (BALT). GALT which include the tonsils, adenoids, and appendix, and specialized structures called Peyer's patches in the small intestine, collect antigen from the epithelial surfaces of the gastrointestinal tract. In Peyer's patches, the most important and highly organized of these tissues, antigen is collected by specialized epithelial cells known as M cells. In Peyer's patches lymphocytes form a follicle consisting of a large central dome of B-cells surrounded by T lymphocytes. Similar but more diffuse aggregates of lymphocytes protect the respiratory epithelium, where they are known as BALT. Lymphoid tissue is also found in the genitourinary tracts. Mucosal immunity is discussed in further detail in the *Introduction* part, section 1.

Although very different in morphology, the lymph nodes, spleen, and mucosal-associated lymphoid tissues all share the same basic architecture. Each of these tissues operates on the same principle; trapping antigen from sites of infection and presenting it to migratory small lymphocytes, thus inducing adaptive immune responses.
The secondary lymphoid tissues also provide signals to the lymphocytes that do not encounter their specific antigen, so that they continue to survive and re-circulate until they encounter an antigen with adequate specificity. This is of immense importance in the maintenance of the correct amount of circulating T- and B- lymphocytes.

It also ensures that only lymphocytes with the potential to respond to foreign antigen are sustained. B- and T-lymphocytes that have matured in the primary lymphoid tissue, but have not yet encountered antigen that match their specificity, are referred to as naive lymphocytes. These cells circulate continually from the blood into the peripheral lymphoid tissues, by squeezing between the cells of capillary walls. They are then returned to the blood via the lymphatic vessels or, in the case of the spleen, returned directly to the blood.

Pathogens can enter the body by numerous routes and set up an infection practically anywhere. However, antigen and lymphocytes will eventually encounter each other in the peripheral lymphoid organs. Circulating T- and B-cells continuously passes through secondary lymphoid tissue, to which antigen is also carried from sites of infection, primarily within macrophages and dendritic cells. In the lymph nodes, these antigen-presenting cells display the antigen to T-cells, which they also help to activate. B-cells that encounter antigen as they tour the lymphatic system are also arrested and activated in the lymph node.
In the event of an infection antigen-specific lymphocytes, that have recognized an infectious agent, are arrested in the secondary lymphoid tissue. There they proliferate and differentiate into effector cells capable of combating the infection.

MHC-molecules; a brief introduction to antigen presenting tools

Majorhistocompatibility-molecules (MHC), also known as human leukocyte-associated (HLA) antigens, are a cluster of cell surface glycoproteins (figure 2). Class I MHC molecules are expressed on virtually all nucleated cells, whereas class II molecules, are normally confined to specialized cells, such as B cells, macrophages, DC's and other antigen-presenting cells.

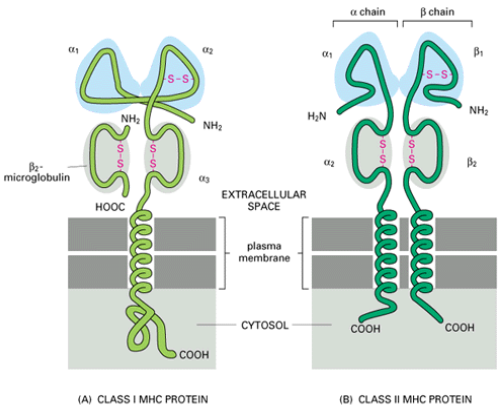


Figure 2. MHC class I and II proteins.
(A, left) Schematic presentation of MHC class I with its three extracellular domains (α1,α2 and α3) non-covalently associated with a β2 microglobuline. While β2 microglobuline is invariant, the α-chain is extremely polymorphic, especially in the α1 and α2 domains.
(B, right) In MHC class II molecules both chains (α and β) are polymorphic. The α1and β1 domains comprise most of the variation. Thus there are striking similarities between MHC class I and class II proteins. In both, the two outermost domains (shaded in blue) interact to form a groove that binds foreign antigen and presents it to T cells.
Figure obtained from; Molecular Biology of the Cell, 3rd ed. Alberts et al. Fig. 23-43.

MHC-class I are membrane bound heterodimers composed of a ~44kD α-chain and a 12kD β2- microglobulin. The α-chain gene encodes three extracellular domains (α_1, α_2 and α_3), a trans-membrane domain and an intra cellular domain that anchors the molecule to the membrane surface. α_1 and α_2 combine to form a groove, in which antigenic peptides can bind. MHC-class I mainly bind and display peptide fragments and proteins that have been synthesized within the host cell such as viral antigens or tumour specific proteins (figure 3).

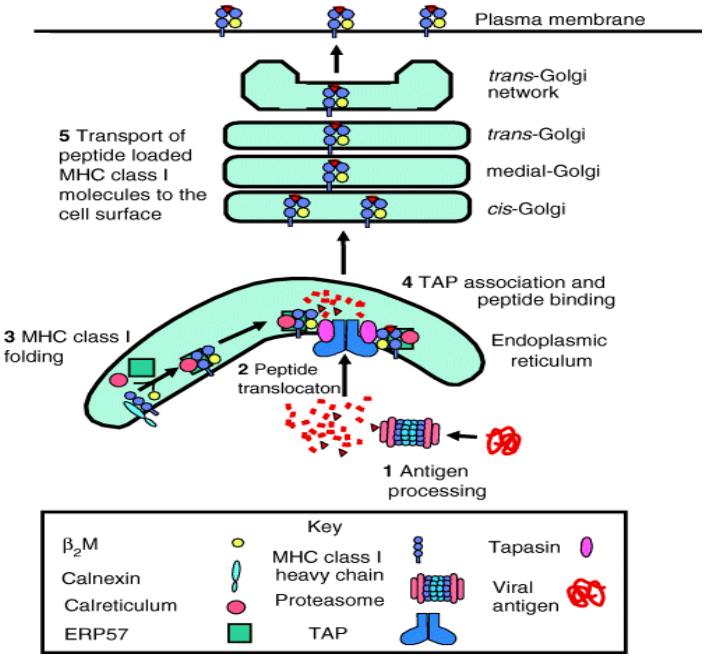


Figure 3.
The MHC class I antigen presentation pathway. (1) Endogenous antigens (i.e. viral proteins, intracellular bacteria, tumour antigens and self-antigens) are proteolytically processed in the cytosol by the proteasome. (2) Peptides generated by the proteasome are translocated into the ER lumen by TAP. (3) MHC class I molecules fold and assemble in the ER lumen with the aid of the ER chaperones calnexin, calreticulum and ERP57. (4) The MHC class I molecule in a complex with calreticulum and ERP57 associates with TAP and tapasin facilitates peptide binding. (5) Peptide loaded MHC class I molecules dissociate from TAP and are transported through the secretory pathway to the plasma membrane. Here, the MHC-I/peptide complexes can be recognized by CD8+ T-cells with complementary TCRs.

Obtained from Hewitt, (2003).

MHC-class II molecules, like class I, are composed of two polypeptide chains. Both chains, designated α (33-35 kD) and β (25-30 kD) are transmembrane proteins. The α-chain has two extracellular domains ($\alpha_1$and α_2) and so does the β-chain (β_1 β_2). The α_1- and β_1 chains combines to form a groove similar to the one found in MHC-class I. MHC-class II primarily binds and display antigens that have been ingested and proteolytically processed, such as bacterial antigens engulfed by phagocytic cells (see figure 4).

The composite structure of self-MHC and antigenic peptide is the target for T-cell receptors (TCRs).

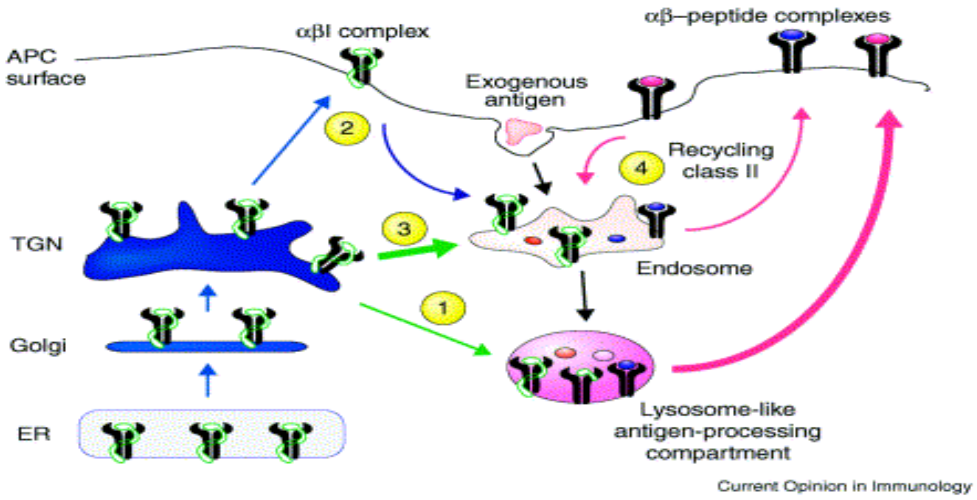


Figure 4. Transport and antigen pressentation of MHC class II molecules
Class II αβ dimers (black pincers) associate with Ii trimers (white snakes) in the ER (bottom left of the figure) and the resulting αβ Ii complexes are transported through the Golgi apparatus. In the TGN, αβ Ii complexes are then diverted from the constitutive secretory pathway into the endocytic pathway. Transport of these complexes to antigen-processing compartments can occur following a direct pathway from the TGN to lysosomes (Pathway 1), indirectly after transit through to the cell surface and endosomes (Pathway 2), or indirectly following a pathway of TGN-to-endosome transport (Pathway 3). Ii is degraded in antigen-processing compartments and class II complexes are loaded with peptide epitopes from exogenous antigens (i.e. bacteria, free viruses, yeasts, protozoa, and toxins) (different peptides are depicted using different colors). Antigenic peptides and αβ–peptide complexes are subsequently sent to the cell surface (thick pink arrow). Pre-existing surface αβ–peptide complexes can also internalize and recycle through endosomes, where peptides can be exchanged (Pathway 4, thin pink arrows). At the plasma membrane, the peptide and MHC-II complexes are presented for CD4+ T-cells with complementary TCRs.

Figure obtained from Hiltbold, (2002).

T-lymphocytes; development, antigen recognition and effector functions

T-cells are defined by expression of the heterodimeric αβ T-cell surface receptor (figure 5) and specific co-receptor expression. TCR α- and β chains are transmembrane polypeptides with extracellular, immunoglobulin like, variable domains (VαVβ) and a constant domains (CαCβ).

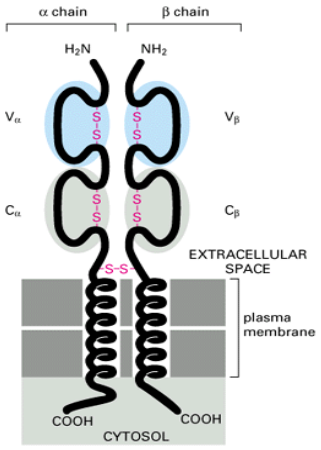


Figure 5. Structure of the TCR
The hetrodimeric T-cell receptor is composed of α and β polypeptide chains. They are transmembrane, glycosylated chains with an extracellular part that is folded into two Ig-like domains, the membrane distal variable (V) and the constant (C) domains.
An interchain disulfide bond links the two polypeptides. The antigen-binding site, that recognizes self-MHC-peptide complexes, is formed by the Vα and Vβ domain (shaded in blue).

Figure obtained from; Molecular Biology of the Cell, 3rd ed. Alberts et al. Figure 23-42.

The variable domains of the TCRs include hyper variable loops, or complementarity-determining regions (CDRs) 1, 2, and 3. Each T-cell bears TCRs with a single specificity. The huge receptor repertoire is enabled by the variation found within the CDRs and rearrangement of the elements therein.

Selection for fully functional T-cells occurs in the thymus where the T-cells mature. T-lymphocytes that undergo positive selection receive survival signals through interaction with self-MHC-peptide complexes and other ligands and are actively rescued from apoptosis. Negative selection refers to the deletion or inactivation of potentially auto reactive T-cells. These cells do not receive protective signals and are sentenced to "death by neglect". Cells that fail any of these, positive and negative, selection criteria are removed by apoptosis. It has been calculated that fewer than 5% of the developing T-lymphocytes escape the well guarded control gates of the thymus to enter circulation (reviewed in Sprent and Kishimoto, 2001).

During their maturation in thymus, T-cells differentiate into discrete subpopulations with defined effector functions. Two major categories are defined by selective expression of CD4 or CD8. CD4+ T-lymphocytes are generally known as T-Helper cells (Th) and CD8+ T-cells are known as cytotoxic T-cells.

As mentioned briefly above TCRs recognize antigens only in the context of self-MHC molecules with antigenic peptide displayed in its groove. The TCR interacts both with the presented peptide and with the flanking α-helixes of the MHC-groove. This concept, known as MHC restriction, allows the T-cells to ignore free, extracellular antigens, thereby enabling them to focus their effort towards infected cells. The MHC/peptide/TCR complex is known as the immunological synapse.

Conventional antigen presentation by MHC molecules is limited to protein antigens. Alternative presentation mechanisms for pathogenic lipid- and polysaccharide antigens have evolved. CD1 molecules, structurally related to MHC, displays antigens to T-cells with the common αβ-TCR or to lymphocytes expressing an alternative γδ-TCR.

CD4+ T-cells only recognize antigen presented on MHC-class II molecules, whereas CD8+ T-cells needs to see the antigen in a MHC-class I context. CD8 and CD4 serve as co-receptors for MHCI/peptide and MHCII/peptide recognition, respectively (Gao et al., 2002, review). The α- and β-chains of the TCR associate with invariant accessory molecules, collectively known as the CD3 complex, that serves as a signal transducer.

Interaction between the TCR/CD3 complex and the MHC/peptide complex, does not provide sufficient signals for proper T-cell activation. Adequate activation requires additional co-stimulatory signals. CD28 on T-cells and CD 80/B7.1 or CD86/B7.2 on the APC provides these signals. Properly activated ITAMs (immunoreceptor tyrosine-based activation motifs) in the cytoplasmic parts of the CD3 complex initiates a cascade of enzymatic reactions, leading to activation of genes that control lymphocyte proliferation, differentiation and survival (Sommers et al., 2000).

Both CD4+ and CD8+ cells differentiate to distinct, effector cells upon activation by antigen. Activation of CD4+ T-helper cells (Th) causes production and release of IL-2. At this stage the cells are designated Th0. The Th-cells continue to respond to the activating signals and goes on to diverge into Th1- and / or Th2-cells, distinguished by the cytokine signature they produce. Th1 cells mainly leave an IL-2, Interferon-γ, TNFβ signature, which induces T-cell proliferation and T-cell cytotoxisity, activate macrophages and induces NK- cytotoxisity. This gives them an important role in cell-mediated immune responses. Th2 cells are characterized by IL-4, IL-5, IL-6, IL-10, IL-13 secretion. They play an important role in B-cell maturation and antibody production (discussed in detail in the next section). Most immune responses will show elements of both Th1 and Th2 differentiation.

Naïve CD8+ T-lymphocytes are activated by peptides carried by self-MHC class I molecules. Most cells presents peptides derived from endogenous proteins on MHC class I molecules. Thus, by doing so, they become targets of the effector function of $CD8^+$ T cells. Professional APCs, primarily DCs, have been demonstrated to load peptides derived from internalized antigens onto MHC class I molecules. This process, known as cross presentation, and the subsequent activation of immune responses, called cross-priming, consequently leads to cytotoxic immune responses against exogenous antigens presented in the context of MHC class I (Amigorena, 2003; Norbury and Sigal, 2003, review).

Activated CD8+ T-lymphocytes proliferate and differentiate to become cytotoxic effector lymphocytes (CTLs). CTLs develop exocytic granules containing perforin and granzymes upon activation. They also produce and secrete cytokines. When re-challenged with antigen, the CTLs direct their lethal load towards the site of target-contact and release their effector molecules to induce target cell death. This pathway is crucially important for the ability of CTLs to induce target cell death in many biologically important scenarios, including allogenic rejection, tumour cell killing, and the clearance of some viruses.

The cytotoxic CD8+ T-cells kills via yet another mechanism, known as the Fas-FasLigand induced apoptosis. Killing of the target cell is mediated by surface molecule interaction between apoptosis triggering Fas on the target cell and Fas-Ligand on the CTL (Russell and Ley, 2002, review).

A key feature of adaptive immunology is the defence's ability to remember the antigens it has previously encountered. This is called immunological memory. Both CD4+ and CD8+ T-cell develop into memory cells. Upon antigen re-challenge they are induced to rapidly synthesize large quantities cytokines and they expand rapidly. Memory cells require lower levels of Ag and co-stimulation than naive T cells to be activated. Moreover, they are recruited more rapidly to sites of inflammation than are naïve T-cells (reviewed in Swain, 2003).

B-lymphocytes; development and immunoglobulins

B-lymphocytes originate from bone marrow stroma stem cells. During IL-7 dependent maturation in the bone marrow, B-cells develop to express membrane bound antigen receptors (BCR). The antigen receptors found on B-cells are immunoglobulins (Igs). Ig molecules are composed of two identical heavy chains and two identical light chains. The amino terminal domains of the heavy and the light chain are designated variable-heavy (V_H) and variable-light (V_L) respectively. The variable domains of Igs, like TCRs, contain three hyper variable regions, also known as complementarity-determining regions (CDRs). Four gene segments; variable (V), diversity (D) (heavy chains only), joining (J) and constant (C) segments are involved in the formation of the receptor repertoire. Rearranged V_L and J_L genes are coupled with C_L genes to produce light chains. Light chains come in to versions known as κ and λ respectively. Heavy chains contain $V_HD_HJ_H$ genes coupled to C_H genes. Additional variation is introduced to the receptor repertoire by inaccurate splicing, frame shift and introduction of additional nucleotides.

Heavy chain constant genes contain exons that encode different constant regions that determine classes and subclasses of immunoglobulins. Combination of V_H and V_L domains produces the antigen binding site on the receptor, thus each Ig contains two identical antigen binding sites. The C-terminal, constant domains of the two heavy chains pairs to form the Fc-domain of the Ig-molecule that is responsible for the effector functions mediated by Igs. Like TCRs, the BCR is dependent on invariant transmembrane proteins for signal transduction. Igα and Igβ (CD79 a, b), both containing ITAM domains, communicate intracellular signals initiated by BCR-antigen interaction.

Naïve, developing B-cells initially express membrane bound receptors with μ-heavy chains known as isotype IgM. Later in lymphocyte development the same $V_HD_HJ_H$ genes are coupled with δ heavy chains to produce isotype IgD receptors. B-cells produce receptors with only one antigen specificity. This is ensured through the process of allelic exclusion. Mature naïve B-cells, expressing IgM and IgD, leaves the bone marrow, enters circulation and migrate to secondary lymphoid tissue. In the lymphoid tissue B-cells encounter antigens.

Antigens are recognized by the membrane bound receptor, internalized, processed and re-expressed on the B-cells MHC class II molecules. T-cells that recognize the B-cell presented MCH-peptide complex may become activated and start to produce cytokines that stimulate the antigen presenting B-cell to proliferate and differentiate to antibody producing plasma cells. This B-cell activation pathway, in which B-cells act as APCs, is known as a T-cell dependent response.

Supplementary T-lymphocyte interaction, particularly the binding of CD40 on B-cells to CD40-ligand on T-cells, induces an important event known as class switching. This process, dependent on the cytokine profile released by the Th-cell, shifts the initial IgM response towards immune responses dominated by antibodies of other isotypes. However, the antigen specificity remains the same as the VJD-genes are not rearranged outside of the bone marrow. B-cell activation occurs mainly in the germinal centres of secondary lymphoid tissue.

As B-cells undergo class switching in the germinal centres, an active process introduces mutations in the variable domains of the light and heavy chains. This may lead to loss of affinity for the triggering antigen, in which case the B-cell looses receptor mediated growth signals necessary for life.

On the other hand, if the somatic mutations lead to increased affinity for the antigen, the antibody producing cell get a propagating advantage, and grow to outnumber the other responding cells. This process is known as affinity maturation. Recent investigations indicate that a protein expressed specifically in activated B cells, activation-induced deaminase (AID), causes deamination of cytosine residues within the immunoglobulin loci, thereby triggering major programs of antibody gene diversification. It is now apparent that programmed introduction of uracil into DNA by targeted cytosine deamination is used by the adaptive immune system as a means of triggering multiple pathways for somatic modification of antibody genes (reviewed in Neuberger et al., 2003).

Plasma cells leave the lymphoid tissue to secret antibodies with specificity corresponding to the specificity of the membrane bound Ig.

Following class switching, some of the activated B-lymphocytes go on to become memory B-cells. These react rapidly with a characteristic IgG response to antigen re-challenge. B-cell memory is of critical importance to the success of vaccination.

B-cells may also be activated independent of T-cells. In the absence of T-cell co-stimulatory protein signals, monomeric antigens are unable to trigger B-cell activation. However, polymeric antigens, such as certain pathogenic polysaccharides, lipids and proteins with repetitive structures, are able to cross-link and cluster membrane IgMs, thereby activate the receptor bearing cell. Class shifting, somatic hyper mutation and generation of memory does not occur in T-cell independent B-cell activation. Such activation ultimately leads to an IgM limited response of poor specificity and of limited duration.

Antibody structure and function; a brief introduction

Antibodies (Abs) are central molecules in the immune system. They play a crucial role in recognition and elimination of infectious agents.

The general structure of an antibody comprises two identical heavy- and two identical light-Ig-chains. They combine to generate a bifunctional molecule, comprising two highly diverse N-terminal antigen-binding sites and conserved effector mediating sites in the C-terminal end (figure 6).

The infinite diversity of binding specificities found in the variable (V_L and V_H) domains is created by imprecise recombination and somatic mutations as described above. The Ab molecule is frequently divided into substructures; V_L and V_H linked to C_L and C_H1 respectively constitute the Fab part, whereas the effector functions of the Abs are mediated by the carboxy-terminal C_H domains, the Fc portion of the Ab. In addition a flexible part of the heavy chain, containing disulfide bridges, is known as the hinge region. Two Fab fragments linked by parts of the hinge is known as $F(ab')_2$. The minimum fragment for antigen binding, V_L and V_H, is known as fragment of variation or just Fv.

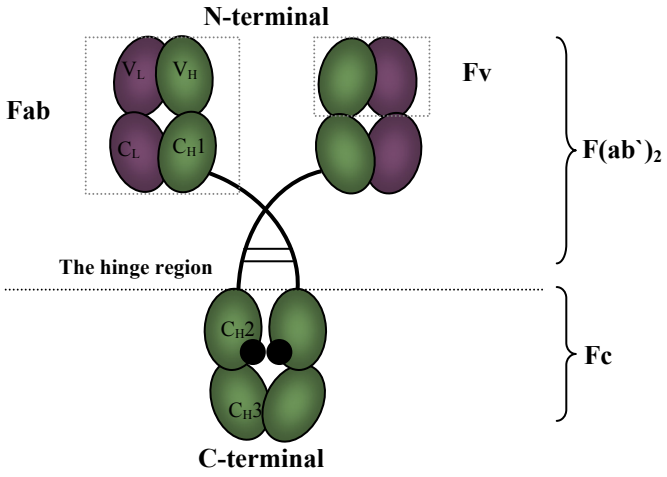


Figure 6. The basic structure of antibodies; represented by IgG
The antibody molecule is made up of two identical light- (purple) and heavy- (green) chains of immunoglobulin domains. The molecule is commonly divided into the following sub structures; The Fc-fragment which consists of CH2 and CH3; The Fv fragment that consists of VH and VL domains; Fab fragments which consist of one light chain (VL-CL) and a part of the heavy chain (VH-CH1) and the hinge region, a flexible region of heavy chain. Note that all domains of IgG pair closely, except for the CH2 domains. The CH2 domains are separated by carbohydrate residues, marked as black spots in this simplified diagram.

Note: The hinge region has been oversimplified and is not to scale with the rest of the molecule.

There are five different human Ab classes, namely IgA, IgD, IgE, IgG and IgM. The characteristics of each class, with focus on human Igs, will be discussed below. As mentioned in the previous section, immunoglobulin classes are distinguished by their heavy chains, denoted by the corresponding lower case Greek letter; α, δ, ε, γ and μ respectively. There are only two different light chains; the λ- and κ chains.

Ig heavy and light chains have common structural domains, called immunoglobulin folds. They are characterised as a sandwich of two extended β-sheets of three and four anti-parallel β-strands in constant domains, and four and five anti-parallel β-strands in the variable domains, linked together by disulphide bridges. The β-strands are connected by inter sheet- and intra sheet loops of variable length and amino acid sequence. Three loops of the variable domain of Abs, CDR1, CDR2 and CDR3, determine the antigen specificity of the antibody.

IgG is the most studied among the Igs and consists of a heavy chain of the structure V_H-C_H1-hinge-C_H2-C_H3 and a light chain of the structure V_L-C_L. The heavy chain and light chain are linked by a disulphide bond between C_H1 and C_L and non-covalent interactions between the V_H/V_L and C_L/C_H1 domains. The heavy chains are linked by disulphide bonds in the hinge region, while the C_H2 domains are separated by carbohydrate residues, and non-covalent interactions exist between the paired C_H3 domains. IgG comes in four versions; IgG1-4 with γ1-4 heavy chains respectively of which IgG1 is dominating in serum.

IgM antibodies have μ heavy chains with four constant Ig-domains. Unlike the membrane bound IgM of B-cells, soluble IgM usually form pentamers. They contain an additional Ig-like protein known as the J-chain, lack the classical hinge region found in IgGs and are heavily glycosylated. If the J-chain is not available, hexamers form. The multimer is held together by inter Cμ3 disulphide bonds and disulphide bonds between the 18 amino acid (aa) tailpiece. The pentamer possesses 10 identical antigen binding sites, the hexamer 12.

IgA is secreted by plasma cells as monomers and as dimers (pIgA). The α- heavy chains, comes in two versions (α1 and α2), contain three constant domains (Cα1-3), are glycosylated and holds a hinge region very much like the one found in IgG. pIgA, like pentameric IgM, associates with J-chain that connects the two monomers at their C-terminal tailpiece, giving the molecule a total of 4 antigen binding sites. IgA plays a major role in mucosal immunity. It is present in secretes as secretory IgA (sIgA) coupled to the secretory component (SC), the extracellular part of the polymeric Ig receptor (pIgR) found on epithelial cells.

IgD constitutes only 1% of the total serum Ig content and is known to serve as BCRs when membrane bound to B-lymphocytes. IgD heavy chains contain three δ constant domains (Cδ1-3), a short tail piece, and a single disulphide bond in the hinge region. The soluble form is known to be susceptible to spontaneous proteolytic degradation.

The fifth Ig class, IgE, have heavy chains with four constant domains (Cε1-4). They lack the classic hinge region seen in IgG. Circulating IgE is scarce. However, it is frequently found bound to the Fcε receptor on basofils, eosinophils and mast cells.

The role of the Ab molecule is to act as an adaptor molecule. Thus, they link target antigens to effector cells and molecules that mediate elimination of the target. The effector functions of Abs can be divided into two categories: Firstly, those mediated by binding complement component C1q, resulting in activation of the classical complement pathway. Secondly, those mediated by binding of Fc-domains to Fc receptors (FcRs) on effector cells.

Fc receptors; antibody docking stations, transporters and maintainers

In addition to being the docking point for Abs on immune effector cells, Fc receptors facilitate specialized transport of antibody molecules and finally, they are involved in Ig homeostasis.

FcRs for IgA, IgG and IgE are present on the surfaces of several cells of the immune system. These receptors FcαRs (FcαRI), FcγRs (FcγRI,II,III) and FcεRs (FcεRI, II) that bind IgA, IgG and IgE respectively, interact with antibody-antigen complexes to activate effector responses in effector cells such as phagocytes, NK cells, mast cells and cytotoxic T-lymphocytes. Binding of Igs to these FcRs triggers biological responses such as phagocytosis and antibody-dependent cell-mediated cytotoxisity (ADCC), mast cell degranulation, release of mediators of inflammation, and the regulation of B cell activation and antibody production.

Neonatal Fc receptor (FcRn) initially discovered to transfer IgG molecules from the mother to the fetus or infant, hence the name, binds and release IgG in a pH dependent manner (high affinity at slightly acidic pH and low affinity at neutral pH) enabling efficient unidirectional transport. FcRn now appears to be used for transport of IgG in sites other than the intestine, and in adults as well as in neonates. FcRn is also known to be involved in IgG homeostasis. The receptor provides the host with a recycling mechanism, leading to increased IgG serum life.

Another Ig transporting receptor, the polymeric Ig receptor (pIgR), mediates transport of polymeric IgA and IgM across epithelial cell layers.

Catalytic activity of antibodies; novel effector functions

Recent work has demonstrated that antibodies have a new effector function independent of complement and FcR mediated antigen elimination. Abs may in fact harbour recognition abilities and killing potential within the same molecule.

All antibodies, regardless of antigenic specificity, can catalyse the generation of highly reactive and bactericidal oxidants including, hydrogen peroxide (H_2O_2), dihydrogen trioxide (H_2O_3) and a component with the chemical signature of ozone (O_3) from singlet oxygen ($^1O_2$) and water (H_2O). Neutrophiles have the potential both to produce $^1O_2$ and to bind Abs, and are likely to be a biological source of O_3 (Babior et al., 2003; Wentworth et al., 2002; Wentworth et al., 2003).

Concluding remarks:

This general introduction to the immune system has been included in my thesis to provide the non-immunologist reader with crucial background knowledge for better understanding of the work presented in the enclosed manuscript. It is beyond the scope of this thesis, and has not been my intention, to write a comprehensive and detailed review of the complex and fascinating biology of the immune system. However, I do hope the reader find my presentation interesting and informative. Interested readers are referred to other excellent reviews (Chaplin, 2003; Parkin and Cohen, 2001) and text books (Janeway, 2001; Roitt, 2001) on the immune system.

References

Alberts B., Bray D., Lewis J., Raff M., Roberts K. and Watson J. D. (1994) *Molecular Biology of the Cell*. Garland Publishing, Inc., New York.

Amigorena S. (2003) Y in X priming. *Nat Immunol* **4**, 1047-8.

Babior B. M., Takeuchi C., Ruedi J., Gutierrez A. and Wentworth P., Jr. (2003) Investigating antibody-catalyzed ozone generation by human neutrophils. *Proc Natl Acad Sci U S A* **100**, 3031-4.

Boyce J. A. (2003) Mast cells: beyond IgE. *J Allergy Clin Immunol* **111**, 24-32; quiz 33.

Chadburn A. (2000) The spleen: anatomy and anatomical function. *Semin Hematol* **37**, 13-21.

Chaplin D. D. (2003) 1. Overview of the immune response. *J Allergy Clin Immunol* **111**, S442-59.

Cole D. S. and Morgan B. P. (2003) Beyond lysis: how complement influences cell fate. *Clin Sci (Lond)* **104**, 455-66.

Gao G. F., Rao Z. and Bell J. I. (2002) Molecular coordination of alphabeta T-cell receptors and coreceptors CD8 and CD4 in their recognition of peptide-MHC ligands. *Trends Immunol* **23**, 408-13.

Hewitt E. W. (2003) The MHC class I antigen presentation pathway: strategies for viral immune evasion. *Immunology* **110**, 163-.

Hiltbold E., M., Roche, P., A. (2002) Trafficking of MHC class II molecules in the late secretory pathway. *Curr Opin Immunol.* **14**, 30-5.

Janeway C. A. T., Paul; Walport, Mark; Shlomchik, Mark. (2001) *Immunobiology*. Garland Publishing, New York and London.

Karre K. (1981) On the immunobiology of natural killer cells. Doctoral thesis, *Karolinska institute*.

Karre K. (2002) NK cells, MHC class I molecules and the missing self. *Scand J Immunol* **55**, 221-8.

Marshall J. S., King C. A. and McCurdy J. D. (2003) Mast cell cytokine and chemokine responses to bacterial and viral infection. *Curr Pharm Des* **9**, 11-24.

Middleton D., Curran M. and Maxwell L. (2002) Natural killer cells and their receptors. *Transpl Immunol* **10**, 147-64.

Neuberger M. S., Harris R. S., Di Noia J. and Petersen-Mahrt S. K. (2003) Immunity through DNA deamination. *Trends Biochem Sci* **28**, 305-12.

Nielsen C. H. and Leslie R. G. (2002) Complement's participation in acquired immunity. *J Leukoc Biol* **72**, 249-61.

Nonaka M. (2001) Evolution of the complement system. *Curr Opin Immunol* **13**, 69-73.

Norbury C. C. and Sigal L. J. (2003) Cross priming or direct priming: is that really the question? *Curr Opin Immunol* **15**, 82-8.

Panaro M. A., Brandonisio O., Acquafredda A., Sisto M. and Mitolo V. (2003) Evidences for iNOS expression and nitric oxide production in the human macrophages. *Curr Drug Targets Immune Endocr Metabol Disord* **3**, 210-21.

Parkin J. and Cohen B. (2001) An overview of the immune system. *Lancet* **357**, 1777-89.

Roitt I., Brostoff, J. and Male, D. (2001) *Immunology*. Harcourt publishers ltd, London.

Russell J. H. and Ley T. J. (2002) Lymphocyte-mediated cytotoxicity. *Annu Rev Immunol* **20**, 323-70.

Sommers C. L., Dejarnette J. B., Huang K., Lee J., El-Khoury D., Shores E. W. and Love P. E. (2000) Function of CD3 epsilon-mediated signals in T cell development. *J Exp Med* **192**, 913-19.

Sprent J. and Kishimoto H. (2001) The thymus and central tolerance. *Philos Trans R Soc Lond B Biol Sci* **356**, 609-16.

Swain S. L. (2003) Regulation of the generation and maintenance of T-cell memory: a direct, default pathway from effectors to memory cells. *Microbes Infect* **5**, 213-9.

Takeda K., Kaisho T. and Akira S. (2003) Toll-like receptors. *Annu Rev Immunol* **21**, 335-76.

Underhill D. M. (2003) Toll-like receptors: networking for success. *Eur J Immunol* **33**, 1767-75.

Upham J. W. (2003) The role of dendritic cells in immune regulation and allergic airway inflammation. *Respirology* **8**, 140-8.

Walport M. J. (2001a) Complement. First of two parts. *N Engl J Med* **344**, 1058-66.

Walport M. J. (2001b) Complement. Second of two parts. *N Engl J Med* **344**, 1140-4.

Wentworth P., Jr., McDunn J. E., Wentworth A. D., Takeuchi C., Nieva J., Jones T., Bautista C., Ruedi J. M., Gutierrez A., Janda K. D., Babior B. M., Eschenmoser A. and Lerner R. A. (2002) Evidence for antibody-catalyzed ozone formation in bacterial killing and inflammation. *Science* **298**, 2195-9.

Wentworth P., Jr., Wentworth A. D., Zhu X., Wilson I. A., Janda K. D., Eschenmoser A. and Lerner R. A. (2003) Evidence for the production of trioxygen species during antibody-catalyzed chemical modification of antigens. *Proc Natl Acad Sci U S A* **100**, 1490-3.

Witko-Sarsat V., Rieu P., Descamps-Latscha B., Lesavre P. and Halbwachs-Mecarelli L. (2000) Neutrophils: molecules, functions and pathophysiological aspects. *Lab Invest* **80**, 617-53.

Antibody engineering

Antibody molecules and derivatives thereof have proved useful in a variety of research, diagnostic, and therapeutic applications. Their ability to recognize and bind specifically to an infinite number of targets makes them valuable tools in the lab and the clinic. According to the US Food and Drug Administration (FDA), engineered antibodies now represent more than 30% of biopharmaceuticals in clinical trials (reviewed in Hudson and Souriau, 2003).

Intact antibodies; potential and challenges

Initially, clinical use of Abs *in vivo* employed monoclonal mouse Abs (mAbs) generated by classical hybridoma technology (Kohler and Milstein, 1975). This often resulted in unwanted human-anti-mouse immune responses due to their immunogenicity. It is evident that the Abs designated for clinical use must be as human as possible. Genetic engineering has allowed for development of chimeric, humanized Abs, constructed from murine variable Ig domains and humane constant domains. Mouse complementarity determining region (CDR) grafted on to a human Ig framework has also been attempted.

More recently it has become achievable to produce purely human antibodies in transgenic animals (Clark, 2000, review). The use of complete Ab molecules as therapeutical agents has some evident advantages. Intact Abs are divalent, they possess two bindings sites for the target of interest, giving them high functional affinity (avidity) for its target. As described earlier, Abs recruits cytotoxic effector mechanisms through its C-terminal Fc domain. Thus intact antibodies comprising Fc domains are powerful tools. On the down side, however, it should be noted that intact mAbs penetrate tissue poorly, they show slow blood clearance resulting in significant, potentially toxic exposure to untargeted tissue (reviewed in Batra et al., 2002). In addition mAbs have proven to be extremely expensive to develop and manufacture (Torphy, 2002). Despite these limitations, FDA approved, intact Ab-molecules prove successful both in diagnostics and therapy. Their mode of action in therapy or diagnostics is to block receptors on target cells or to operate as conjugates of therapeutic payloads or diagnostic reagents (reviewed in Hudson and Souriau, 2003).

Genetically modified Ab fragments; harmless target binders?

Regardless of the shortcomings of mAbs in therapeutics, antibody based pharmaceuticals still show an enormous potential (Souriau and Hudson, 2003, review). This has been the driving force in development of alternative antibody based technology.

The domain structure of Abs permits their division into functional subunits, or modules, which can be mixed and matched to create novel molecules with a specific subset of functional properties. In some clinical applications, such as receptor blocking, cytokine inactivation or viral neutralization, Fc domains are unwanted and can be removed by uncomplicated proteolysis of intact antibodies to yield Fab or $F(ab')_2$ fragments (see figure 6 and 7). These antibody subcomponents show increased tissue permeability and higher clearance rate in normal tissue than intact Abs from which they originate.

scFvs, constructs of V_L and V_H Ig domains (see figure 7) genetically linked together by a short, flexible, stabilizing inter-domain peptide chain, show even better tissue permeability than Fab fragments (reviewed in Batra et al., 2002). Molecular cloning and bacterial expression of antibody fragments, including the variable region genes of Ig, have greatly facilitated the generation of engineered antibodies. Fab and scFv fragments can be readily genetically engineered, selected against a target of interest in protein display systems, e.g. phage display, and expressed in inexpensive prokaryotic systems (reviewed in Kretzschmar and von Ruden, 2002). Fabs and scFvs are monovalent molecules with reduced functional affinity for their target compared to a complete antibody with the same specificity. This can be overcome by *in vitro* affinity maturation and/or by construction of multivalent molecules, to yield excellent bio-distributable targeting fragments (Batra et al., 2002, review).

Fabs and scFvs do not harbour any cytotoxic effector mechanisms, neither can they recruit such. Fc domains, of intact Abs, depend on proper glycocylation to meet the structural requirements for effector mechanism activation, thus they can only be synthesised in mammalian cells (Krapp et al., 2003). Like mAbs, monospecific target binding Ab fragments may serve as receptor blocking agents or as carriers in toxin-, drug-, prodrug, enzyme-, radionuclide- or dye conjugates.

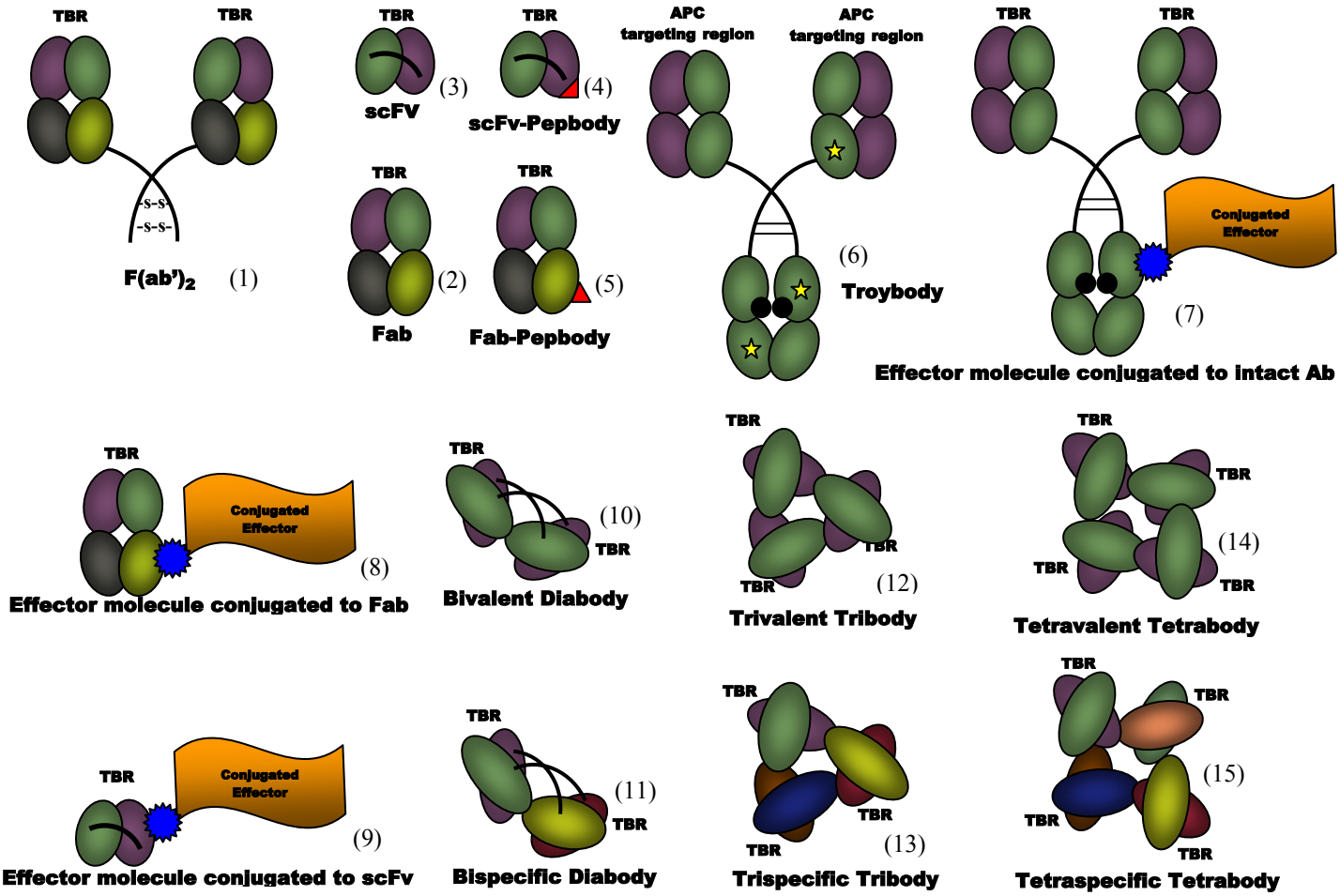


Figure 7. Selection of genetically engineered antibodies and antibody fragments

- Antibody fragments generated by proteolytic cleavage shown as bivalent F(ab')$_2$ (1) and monovalent Fab (2).
- scFv (3); VH and VL linked together by a short polypeptide chain indicated with a black line.
- scFv- and Fab pepbodies (4) and (5) respectively, bind antigen monovalently through its target binding region (TBR) and recruit effector functions via the genetically fused peptide (indicated with red triangle) they display.
- Troybodies (6), a novel vaccination strategy tool, are Abs with one or more T-cell epitopes inserted in the constant domains designed for efficient epitope delivery to APC for which they have specificity.
- Antibodies (7) or Ab fragments (8) and (9) with specificity for a target of interest can be chemically conjugated to effector molecules such as toxins, pro-drugs, enzymes or radionuclides.
- Single chain constructs of bivalent- (10) and bispecific- (11) diabodies, trivalent- (12) and trispecific- (13) triabodies and tetravalent- (14) and tetra specific- (15) tetrabodies. Colours denote multivalens and multispecificity.

Functional antibody fragments: target recognition and elimination

Bispesific antibody fragments, known as diabodies, linking target recognition (i.e. tumour specific antigens) with effector mechanisms (i.e. complement components and cytotoxic cells) have been described (Holliger et al., 1996; Kontermann et al., 1997). Diabodies (see figure 7) initially comprised two different polypeptide chains that needed to undergo proper folding. This called for labours refolding and purification steps.

More recently this hurdle has been overcome by design of single chain constructs of diabodies (dimers), triabodies (trimers) and tetrabodies (tetramers) with multi valence and/or multi specificity (reviewed in Kortt et al., 2001; Todorovska et al., 2001) (figure 7). Mono-chain Ab fragments, known as pepbodies, that possess both antigen binding capabilities and effector functions have recently been described. Pepbodies are Fabs or scFvs fused to one or more peptides that mimic Fc function by binding one or more effector molecules of the immune system (see figure 7). Peptides that bind C1q, pIgR, FcγRs, including FcRn have been selected by means of phage display technology and genetically fused to scFvs and Fabs (Lunde et al., 2002). This promising invention combines advantageous pharmacokinetics, prokaryotic expression, antigen recognition and binding, and engagement of the host's cytotoxic mechanisms in one minimally immunogenic molecule.

Engineered Abs in vaccine development: right on target with Troybodies

Genetically engineered antibodies are also used in the development of innovative vaccine applications. Critical in vaccine development is the design of immunogens that give a strong and specific T-lymphocyte response. CD4+ T-helper cells are central in adaptive immune responses. Consequently, strategies that modulate them are of great interest. Troybodies are engineered antibodies, containing cryptic T-cell epitopes inserted in the loops of their constant domains, with V-genes encoding specificity for antigen presenting cells (Figure 7). Troybodies increase the number of T-cell epitopes that reaches the APC, subsequently increasing the number of peptide-MHC complexes on the APC that ultimately goes on to activate T-helper cells. This new vaccine strategy is flexible and can be designed to target different APCs and induce immunity to numerous T-cell epitopes (Lunde et al., 2001).

Concluding remarks:

This section aims to present the reader with some of the possibilities and prospects of designing, engineering and producing antibody derived drugs, diagnostics and research tools. Furthermore this section elucidates the promises of combining knowledge about the structure and function of antibodies with the powerful tools of molecular engineering in modern medicine. Current technologies employing *in vitro* or *in vivo* selection of fully human antibodies, either by combinatorial approaches or by transgenic animals are review by Brekke and Løset (2003). Their recent review also recapitulates recent developments in antibody discovery, design and therapeutic applications.

References

Batra S. K., Jain M., Wittel U. A., Chauhan S. C. and Colcher D. (2002) Pharmacokinetics and biodistribution of genetically engineered antibodies. *Curr Opin Biotechnol* **13**, 603-8.

Brekke O. H. and Løset G. Å. (2003) New technologies in therapeutic antibody development. *Curr Opin Pharmacol.* **3**, 544-550.

Clark M. (2000) Antibody humanization: a case of the 'Emperor's new clothes'? *Immunol Today* **21**, 397-402.

Holliger P., Brissinck J., Williams R. L., Thielemans K. and Winter G. (1996) Specific killing of lymphoma cells by cytotoxic T-cells mediated by a bispecific diabody. *Protein Eng* **9**, 299-305.

Hudson P. J. and Souriau C. (2003) Engineered antibodies. *Nat Med* **9**, 129-34.

Kohler G. and Milstein C. (1975) Continuous cultures of fused cells secreting antibody of predefined specificity. *Nature* **256**, 495-497.

Kontermann R. E., Wing M. G. and Winter G. (1997) Complement recruitment using bispecific diabodies. *Nat Biotechnol* **15**, 629-31.

Kortt A. A., Dolezal O., Power B. E. and Hudson P. J. (2001) Dimeric and trimeric antibodies: high avidity scFvs for cancer targeting. *Biomol Eng* **18**, 95-108.

Krapp S., Mimura Y., Jefferis R., Huber R. and Sondermann P. (2003) Structural analysis of human IgG-Fc glycoforms reveals a correlation between glycosylation and structural integrity. *J Mol Biol* **325**, 979-89.

Kretzschmar T. and von Ruden T. (2002) Antibody discovery: phage display. *Curr Opin Biotechnol* **13**, 598-602.

Lunde E., Lauvrak V., Rasmussen I. B., Schjetne K. W., Thompson K. M., Michaelsen T. E., Brekke O. H., Sollid L. M., Bogen B. and Sandlie I. (2002) Troybodies and pepbodies. *Biochem Soc Trans* **30**, 500-6.

Lunde E., Rasmussen I. B., Eidem J. K., Gregers T. F., Western K. H., Bogen B. and Sandlie I. (2001) 'Troy-bodies': antibodies as vector proteins for T cell epitopes. *Biomol Eng* **18**, 109-16.

Souriau C. and Hudson P. J. (2003) Recombinant antibodies for cancer diagnosis and therapy. *Expert Opin Biol Ther* **3**, 305-18.

Todorovska A., Roovers R. C., Dolezal O., Kortt A. A., Hoogenboom H. R. and Hudson P. J. (2001) Design and application of diabodies, triabodies and tetrabodies for cancer targeting. *J Immunol Methods* **248**, 47-66.

Torphy T. J. (2002) Pharmaceutical biotechnology, Monoclonal antibodies: boundless potential, daunting challanges. *Curr Opin Biotechnol* **13**, 589-591.

Introduction

1. Mucosal immunity

Mucosal surfaces cover approximately 400m^2 of exposed body surface, more than 200 times of that covered by skin. This vast and vulnerable surface barrier, lining the digestive system, the airways and the genitourinary tracts, is protected by innate mechanisms in intimate cooperation with adaptive mucosal immunity. The mucosa is directly exposed to the external environment and is under constant bombardment of antigens such as bacteria, viruses, food antigens, bacterial toxins and inhaled potentially pathogenic particles. Because mucosal surfaces are ideal portals of entry to most pathogens, mucosal immunity has evolved as a discrete system that performs highly regulated novel immunologic tasks. Mucosa associated lymphoid tissue must continually maintain a delicate balance between active immunity, oral tolerance, and suppression of immune responses (MacDonald, 2003, review).

1.1 Innate mucosal immunity; physical and chemical barriers

Mucosa associated innate immune defence includes physical and chemical barriers. Digestive enzymes of the gastrointestinal tract degrade pathogens by proteolytic cleavage to produce short non-immunogenic peptides. Furthermore, enzymatic alteration of carbohydrate residues on pathogens may lead to loss of virulence. The hostile, acidic environment of the gut provides the host with further protection. A key component of the physical barrier is the mucine gene products. Mucine glycoproteins, secreted by goblet cells, line the apical surface of epithelial cells. This sticky, protective layer prevents pathogens in gaining access to the underlying epithelial cells. Captured potential pathogens are removed with the mucus layer through peristaltic- or mucocilary transport in the gastrointestinal tract and respiratory tract respectively. This process, known as non-immune exclusion, prevents the epithelial cell layer from colonization and invasion. The epithelial cells layer itself is virtually impermeable to macromolecules due to tight junctions both basolaterally and apically. The normal gut flora may provide additional protective activity trough competition against pathogenic bacteria for space and nutrients, thereby preventing colonization and invasion of the epithelia (reviewed in Mayer, 2003).

1.2 Adaptive mucosal immunity

Adaptive immunological protection of mucous membranes is provided mainly by secretory antibodies, predominantly IgA. B-cell responses that give rise to secretory antibodies are triggered mainly in organized lymphoid, epithelial structures where antigens are sampled from the mucosal surface. Gut-associated lymphoid tissue (GALT) constitutes the major part of organized mucosa-associated lymphoid tissue (MALT). Peyer's patches, specialized lymph nodes of GALT, are characterized by follicle-associated epithelium with gateway (M) cells capable of capturing and transporting antigens from the lumen into the underlying lymphoid tissue.

Harmless proteins, components of commensal bacteria or digestive antigens do not only stimulate the secretory antibody system, but also activate suppressive mechanisms collectively known as oral tolerance when induced via the gut. This hypo-responsiveness possibly explains why individuals with intact immune systems show no adverse immune reactions to persistent contact with harmless environmental and dietary proteins. The mechanism governing oral tolerance is not completely understood. However, it is evident that antigen presentation in MALT is of importance. In the absence of inflammation, presentation of peptide to T cells by MHC molecules on APCs, occur in the absence of co-stimulation.

Pathogen derived peptides, on the other hand, concomitantly induce inflammatory responses in the tissue that stimulate the maturation and expression of co-stimulatory molecules on the APC, ultimately leading to a Th1 protective response. B-cells with potential for J-chain expression, that have been activated in MALT (predominantly in Peyer's patches) to produce secretory Abs, migrate through lymph and blood to other mucous tissues (i.e. nasal, salivary, bronchial surfaces, mammary glands, and the urogenital tract) where they are subjected to terminal differentiation and become immunoglobulin secreting plasma cells (Brandtzaeg, 2003, review).

1.2.1 Secretory immunoglobulins; J-chain in polymerization and receptor binding

Approximately 80% of all antibody-secreting plasma cells of the body are located in the gastrointestinal and respiratory mucosa. Some 90% of them are dedicated to production of immunoglobulin A. IgA and IgM antibodies have polymerization signals in the carboxy-terminal constant heavy domains C3α and C4μ respectively (Braathen et al., 2002).

Plasma cells in the lamina propria co-express a 15 kDa polypeptide known as the joining (J) chain. J-chain expression significantly facilitates generation of polymeric IgA (pIgA, primarily dimers) and pentameric IgM (figure 8). J-chain expression may be considered a marker for early B-cells derived from MALT as IgD and IgG secreting tissues are positive for this peptide. The J-chain is exclusively incorporated into IgA and IgM multimers. J-chains do not appear to have a biological function in plasma cells expressing other isotypes of antibodies, thus it is intracellularly degraded.

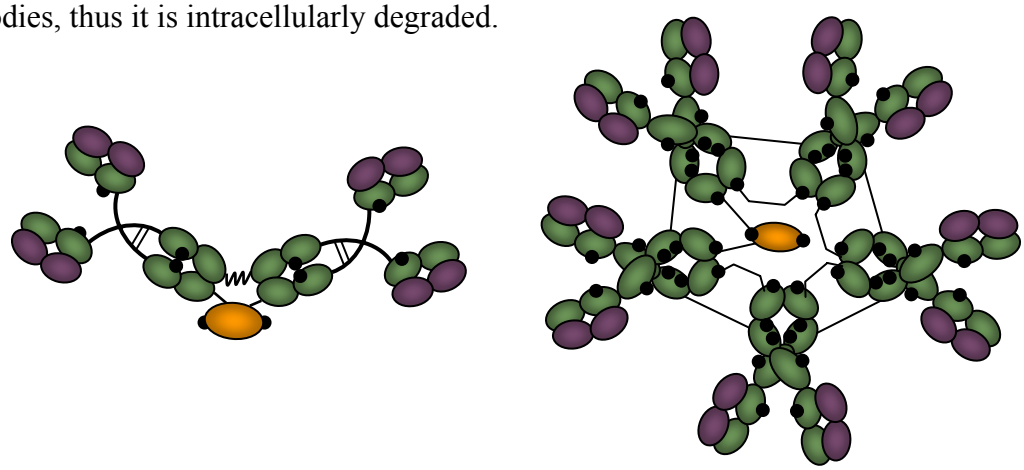


Figure 8. Model of polymeric IgA and IgM containing a J-chain
Left: Polymeric IgAs (pIgA) predominantly form dimmers when a J-chain (orange) is co-expressed.
The glycocylated (black spots) tetra valent complex is the dominating species of Abs in humoral, mucosal immunity.
Right: polymeric IgM form decavalent pentamers bound to one J-chain.
Black lines represent inter-domain disulphide bonds.

Based on figure 4.6, Immunology, 6th ed. Roitt et al.

Polymerization leads to formation of multivalent Ig-complexes with high functional affinity (avidity) for its antigen, making them appropriate tools for agglutinating bacteria, viruses and pathogenic particles. pIgA have little or no complement-activating potential, allowing it function in a non-inflammatory fashion. Furthermore, J-chain-containing pIgs show high affinity for the polymeric Ig receptor (pIgR). This epithelial glycoprotein mediates transfer of dimeric IgA and pentameric IgM across the epithelial monolayer to exocrine secretions (Johansen et al., 2000).

The extracellular domain of the pIgR, known as the secretory component (SC), remains covalently associated with pIgA in secretions. This complex is known as secretory IgA (sIgA). Pentameric IgM also remain associated with SC, but in a non-covalent manner (sIgM). The polymeric Ig receptor is a prerequisite for the protective potential of pIgA and pentameric IgM. The role of antibodies in mucosal immunity is discussed in section 1.2.2.

Although antibodies of all classes may be found in secretions, they are generally not considered as secretory Igs. IgG may play a role in immune exclusion in the respiratory tract. IgG is likely to be subjected to proteolytic degradation in the harsh environment of the digestive system. sIgA and sIgM, on the other hand, show increased stability due to their multimeric structure and their association to SC. Recently, FcRn has been suggested to be responsible for transepithelial transport of IgG to mucosal surfaces (Spiekermann et al., 2002). IgGs role in mucosal immunity remains blurred as it activates complement and leads to inflammation that ultimately causes increased mucosal permeability and consequently, potential antigen penetration. IgE antibodies may be carried by mast cells to mucous membranes and secretions where they cause degranulation and histamine release. Likewise, traces of IgD may be found in secretions, but their biologic significance remains uncertain.

1.2.2 The polymeric Ig receptor: well beyond pIg transport

The polymeric immunoglobulin receptor (pIgR), also known as the transmembrane secretory component, is an epithelial, ~100 kDa glycoprotein of the immunoglobulin supergene family. It consists of five Ig like extracellular domains, an extracellular tail with sites susceptible to proteolytic cleavage and a cytoplasmic tail (figure 9).

Newly synthesized pIgR is transported to the basolateral surface of epithelial cells. pIgRs are endocytosed constitutively and sorted into vesicles for transport to the apical surface or recycled to the basolateral surface. The cytoplasmic domain of the receptor contains the sorting signals that orchestrate endocytosis, transportation and ultimately, exocytosis at the apical (luminal) surface of the epithelial monolayer. This essential, receptor mediated, transepithelial transport mechanism, known as transcytosis, is shared by pIgA and pentameric IgM. The J-chain of pIgA and pentameric IgM constitutes an essential part of the pIgR binding site of antibodies destined for excretion (reviewed in Kaetzel, 2001).

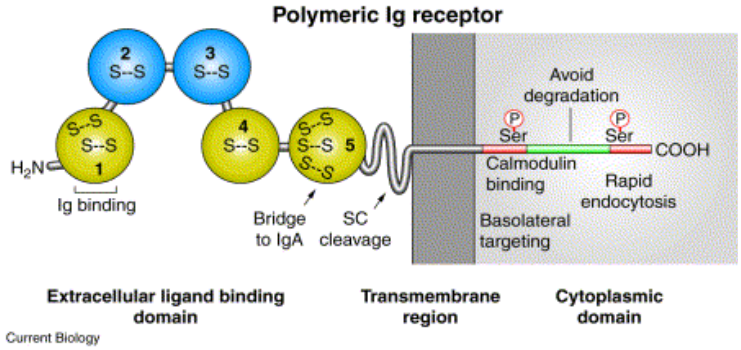


Figure 9. Structure of the pIgR
pIgR is a transmembrane receptor with five extracellular Ig homologous domains and a long cytoplasmic tail containing signals for intracellular sorting and endocytosis. Extracellular domains 1, 4 and 5 (yellow) are highly conserved among species, and contain signals for non-covalent and disulfide bonding to polymeric IgA. Domains 2 and 3 (blue) are less conserved, and are absent in some splice variants of rabbit pIgR. Cleavage of full-length pIgR to release the soluble secretory component (SC) occurs within the linker connecting domain 5 to the transmembrane region.
Obtained from Kaetzel, (2001).

In the further discussion on the role of Abs in mucosal immunity only sIgA will be dealt with in detail, since sIgA is by far the dominating Ab class found in conjunction with mucosal tissue.

The pIgR is instrumental to mucosal immunity. The receptor is involved in multiple mechanisms providing the host with a multi layer defence at the mucosae (figure 10). Receptor-ligand interaction provides two signals sufficient for increased internalization rate and directional transport (Luton and Mostov, 1999). At the apical side of the cell, the 80 kDa extracellular domain of the receptor is proteolytically cleaved off and released as the secretory component (SC) bound to IgA (Brandtzaeg, 1974). Thus pIgR devise the local formation of secretory Ig at the mucosae (figure 10, route 1).

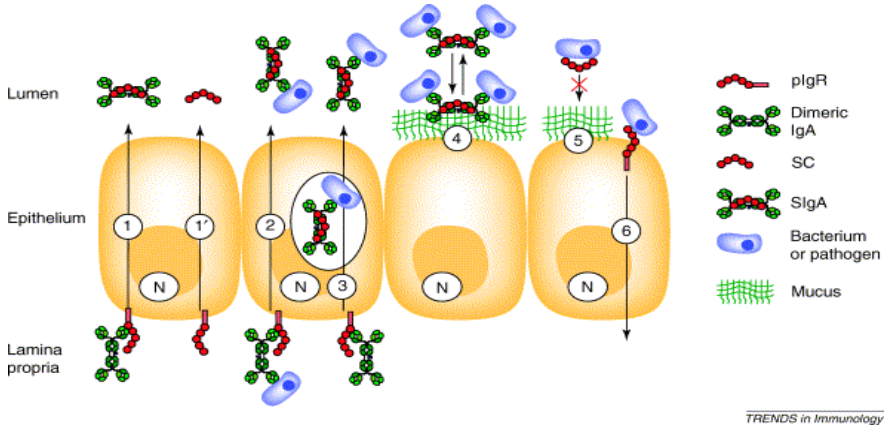


Figure 10. Functions of the polymeric Ig receptor (pIgR) and secretory component (SC)
Route (1): pIgR ensures transcytosis of polymeric IgA (pIgA), and IgM if present, across the epithelial cell layer, leading to the local production of secretory IgA and IgM denoted sIgA and sIgM respectively. sIgM is not shown in this simplified diagram. Route (1'): Transcytosis of unloaded pIgR and release of free SC occurs similarly to route (1), yet with a lesser efficiency. Route (2): pIgR is able to transcytose pIgA loaded with antigens present in the lamina propria suggesting excretion of pathogen-sIgA complexes. Route (3): Abs, en route for luminal secretion, can neutralize invading pathogens by the process of intracellular neutralization. Route (4): Immune exclusion, potentially sIgA's most prominent role, implies neutralization within the lumen, with epithelial anchoring of sIgA mediated by bound SC. Route (5): Free SC might exhibit scavenger function by preventing epithelial cell-pathogen interaction. Route (6): Apical recycling of pIgR may provide pathogens with a cell entrance mechanism. Figure obtained form Phalipon and Corthesy, (2003).

Interestingly, unoccupied pIgR may be internalized on the basolateral surface of the epithelia and released to the lumen by proteolytic cleavage in the same manner as sIgA and sIgM, to form free SC (fSC). Although initially believed to function solely by providing protection against proteolytic degradation it is now evident that this cleavage product, identical to Ab-bound SC, occurs in most secretions where it may contribute to innate mucosal defence (figure 10, route 1'). Free SC has been shown to work as a non-specific scavenger, preventing interaction between luminal microbes or toxins and the epithelia (figure 10, route 5) (de Oliveira et al., 2001).

Dimeric IgA, generated in the lamina propria, may encounter antigens prior to transepithelial transfer. Antigen binding to pIgA and subsequent pIgR mediated transcytosis, consequently leads to removal of the antigen. This excretory pathway (figure 10, route 2) illustrates the protective role of pIgA and pIgR in the tissue underlying the epithelial cell layer. Furthermore, sIgA may come across intracellular antigens, such as some viruses, en route to the apical surface. sIgA-antigen complexes generated intracellularly are excreted to the lumen, thus preventing damage to the epithelium (figure 10, route 3). These two mechanisms can be considered as backup defensive measurements that take care of pathogens that manage to penetrate the epithelial barrier.

sIgA, trapped within the mucus, as antiseptic paint, perform immune exclusion. sIgA is directed and anchored to the mucus by glycosylated SC, where it protects the mucosal barriers from colonization and invasion by binding pathogens (figure 10, route 4) (reviewed in Phalipon and Corthesy, 2003).

Streptococcus pneumoniae is a bacterial pathogen that causes invasive life-threatening disease worldwide. This organism also commonly colonizes the upper respiratory epithelium in an asymptomatic fashion. To invade, this pathogen must traverse the respiratory epithelial barrier, allowing it to cause disease locally or disseminate haematogenously throughout the body. SpsA (also known as CbpA), a surface displayed protein virulence factor of *S. pneumonia* have been shown to bind human pIgR, fSC and sIgA (Hammerschmidt et al., 1997; Hammerschmidt et al., 2000; Zhang et al., 2000). This may turn out to be, for both the pathogen and host, a double-edged sword. Uncleaved pIgR, recycling at the apical face of the epithelia, provides a mechanism for colonization and invasion of the epithelia (figure 10, route 6) (Zhang et al., 2000). Furthermore, pIgR mediated apical to basolateral transport of *S. pneumonia* has been demonstrated in a model system (Zhang et al., 2000). However, this retrograde pIgR transport is disputed and is argued to be bacterial strain- and cell type dependent (Brock et al., 2002).

Normally, pIgR-mediated internalization of microbes is prevented by cleavage of pIgR from the apical surface, the presence of fSC and finally sIgA antibodies in the secretions. However, an imbalance in any of these factors, or an overwhelming infection, could tip the scales in favour of the pathogen (reviewed in Kaetzel, 2001).

Thus, the key components of adaptive mucosal immunity (i.e. pIgR, sAbs and fSC) are able to counteract a variety of strategies employed by invading pathogens, consequently promoting crucial protection at the multiple layers of the epithelial barrier (lumen, epithelium and lamina propria).

Targeting of pIgR as a mechanism for specific delivery of novel therapeutics and diagnostics to epithelial cells and mucosal surfaces is discussed in section 3.

2. Filamentous bacteriophage and peptide display libraries

Phage display is a technology suitable for elucidating protein interactions, such as receptor-ligand interactions, studying molecular evolution and for selection of novel peptides such as therapeutic candidates. To understand the molecular mechanism of protein interactions and to isolate compounds of therapeutic value, it is essential to identify and isolate the genes encoding the proteins of interest. Phage display is a technique by which foreign peptides are expressed on the surface of bacteriophage particles (Smith, 1985). Thus, the viral particles are vehicles that carry within them the nucleotide sequence encoding the surface expressed peptide, and also have the capability to replicate in bacterial hosts. The physical link between genotype and phenotype, provided by the phage particle, is the hallmark of all phage display technology derivatives.

A phage display library is a heterogeneous mixture of phage clones, carrying foreign DNA insertions, thus they display different peptides on their surface. Libraries of vast molecular diversity coding for e.g. peptides or Ab-fragments, expressed as fusions to the coat proteins of the bacteriophage, have been constructed. These libraries, with sequence diversities in the million and billion range, can be conveniently screened by affinity selection, thereby isolating clones with desired properties. In addition to screen naturally occurring protein libraries for ligands, libraries may be constructed *in vitro* to develop peptides with random amino acid sequences and novel activities (Willats, 2002, review).

2.1 *Filamentous bacteriophage: structure and biology*

A number of different bacteriophages have successfully been used in phage display; including F-pilus dependent filamentous phages (Ff), λ, baculovirus, T4 and T7. Members of the Ff phage family, such as the virtually identical M13, fd and fl, are all excellent cloning vehicles due to their tolerance of relatively large DNA insertions. Foreign DNA sequences inserted into their genome is simply accommodated by the assembly of larger phage particles. In the further discussion, filamentous fd-phages will be the focus as they are used experimentally in the work described in this thesis.

Filamentous phage (approximately 1μm in length and 6 nm wide, variation between strains exists) consists of single stranded DNA (ssDNA) packaged into a coat consisting of thousands of copies of the major coat protein gpVIII. Additional minor coat proteins (gpIII, gpVI, gpVII and gpIX) are located at low copy numbers at the ends of the virus (figure 11) (reviewed in Marvin, 1998). All of these surface proteins may serve as fusion partners for expressed peptides (Sidhu, 2001, review). However, most libraries based on filamentous phage are constructed as N-terminal gpIII or gpVIII fusions.

The high copy number of gpVIII on a single phage particle offers the opportunity for multivalent display. However, peptides larger than 6-8 amino acids are not tolerated very well on gpVIII (Iannolo et al., 1995). gpIII, expressed in 3-5 copies, accept larger fusions and have been preferred for display of larger proteins like Ab variable domains or Fab fragments. gpIII display may also provide multivalent display.

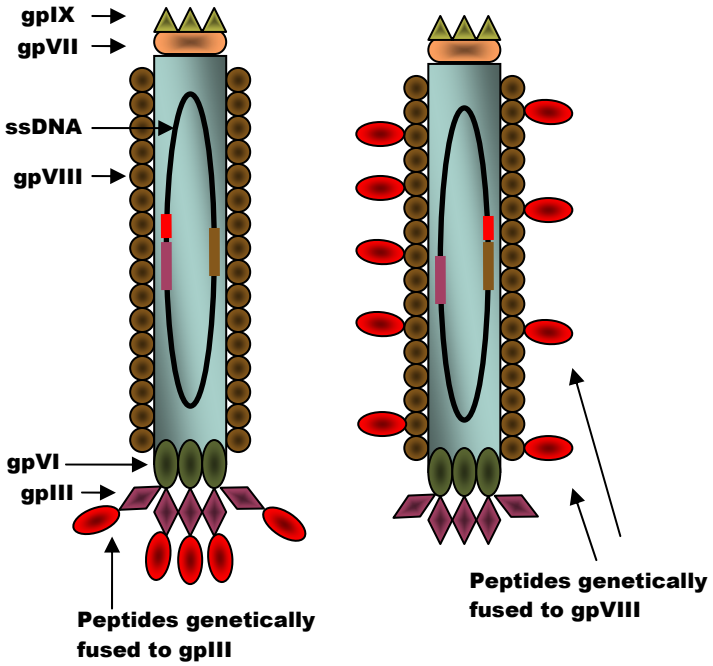


Figure 11.
Schematic illustration of phage display
Filamentous bacteriophage ssDNA (black oval line) is enclosed by thousands of copies of the major coat protein gpVIII (brown), making up most of the virus capsule. The ends of the phage are constructed from, low copy number, minor coat proteins gpVII (orange), gpIX (green triangles), gpVI (dark green ovals) and gpIII (purple).

Left; Display of introduced peptides (red) genetically fused to gpIII as indicated by corresponding colours in the phage genome (purple and red).

Right; Display of a large number of introduced peptides (red) genetically fused to gpVIII as indicated by corresponding colours in the phage genome (brown and red).

Multicopy display consequently leads to an increase in the avidity of the phages for the target they are panned against, making it difficult to discriminate between phages with minor differences in affinity. When selecting high affinity peptides, monovalent display may be advantageous (Lowman et al., 1991). Use of a phagemide vectors may result in monovalent display. Phagemids, plasmids with phage origin of replication, carry sequences encoding the fusion peptides and lack other vital phage genes.

The majority of the bacteriophage genes are provided by a helper phage that is co-infected together with the phagemide into the bacterial host. Thus, secreted viral particles will display a mixture of fusion and wild-type protein, resulting in an average of less than one copy of fusion protein on the phage surface.

The N-terminal gpIII domain plays a crucial role in host infection. First, the phage attaches to the F pilus of male (F+) *Escherichia coli* (*E.coli*) strain as the primary receptor.
Next, the phage attaches to the carboxy-terminal domain of the periplasmic *E.coli* protein TolA, the co-receptor for phage infection (Riechmann and Holliger, 1997).
Subsequently, the viral ssDNA enters the cytoplasm. In the bacterial cytoplasm, the ssDNA plus strand is converted by the host's enzymatic machinery into a double-stranded replication form (RF) by the rolling circle mechanism. The non-lytic propagation mechanism of Ff phage requires the newly synthesised ssDNA and coat proteins to be transported through the bacterial inner membrane prior to assembly and extrusion. Other phages, including λ and T7, assemble to form new virus particles entirely within the cytoplasm of their host causing lysis at time of escape (reviewed in Marvin, 1998; Willats, 2002).

2.2 Phage display in the fUSE5 system

fUSE5 is a phage display vehicle and library system based on filamentous bacteriophage strain fd-tet (Zacher et al., 1980). The fUSE5 vector, constructed for gpIII display, carries a tetracycline resistance gene inserted in a manner that renders the phage defective in minus strand synthesis resulting in low levels of RF DNA and low phage copy numbers. This feature is advantageous as it prevents the host from being killed, allowing for the possibility of growing the peptide presenting phage and host bacteria for an extended period of time. fUSE5 carries a frame shift mutation in gene III, rendering the phages non-infectious.

Replacement of the fUSE5 stuffer fragment with a suitable insert, on *Sfi*I restriction sites, restores the reading frame and makes the phage infectious. There are two *Sfi*I restriction sites in fUSE5, which have non-identical, non-complementary 3-base 3´ overhangs allowing for directional cloning (Scott and Smith, 1990).

2.2.1 fUSE5 random peptide libraries

Random peptide libraries may be constructed by inserting synthetic oligonucleotides with random sequences into the gene encoding the phage coat protein (Scott and Smith, 1990). Peptides can be displayed in fUSE5 as linear fusions or constrained (cyclic) fusions to gpIII. Constrained peptide libraries are constructed by introduction of two cystein residues flanking the random peptide sequences. The cystein residues are capable of forming a disulphide-bridge, reducing the conformational flexibility of the peptide. Thus, they are expected to facilitate formation of tertiary structures that leads to enhanced binding affinity compared to linear peptides (O'Neil et al., 1992). The diversity of the library is limited by a number of factors. The non-lytic biology of the fUSE5 phage allows display of peptides that tolerate the viral assembly at the host's membrane only. Furthermore, peptides that show host cytotoxisity are likely to be eliminated from the library. Moreover, peptides sensitive to proteases found within the host bacteria are eradicated, thus reducing the diversity of the library displayed on the phage particles. In addition the proteolytic activity of the host cell, may lead to reduced valency. That is, reduction of the number of displayed peptides per phage particle.

2.3 Affinity selection; finding the needle in a molecular haystack

Random phage display peptide libraries frequently contain as many as 10^9 different clones. Libraries typically hold hundreds of infective copies of each clone, thus the total number of functional phage in a library may exceed 10^{11} particles. The enormous diversity in the primary library population must be reduced to a manageable level through efficient screening procedures prior to detailed analysis of selected clones. Affinity selection, also known as bio-panning, is the general approach to phage peptide display library screening (figure 12).

The simplest and most widely used approach is to immobilize a target of interest (e.g. your favourite receptor) to a supportive surface (i.e. tubes, wells, beads or membranes) and then expose the library to the target (see figure 12, left panel).

Panning may also be performed directly against cell surfaces, tissues and even animals. Following interaction between the target and library, non-binders are removed by washing (see figure 12, centre panel). Bound phage clones are subsequently eluted by disruption of the target-peptide interaction (e.g. by dramatic changes in pH). Next, eluted binders are propagated by infecting fresh bacterial host cells (see figure 12, right panel).

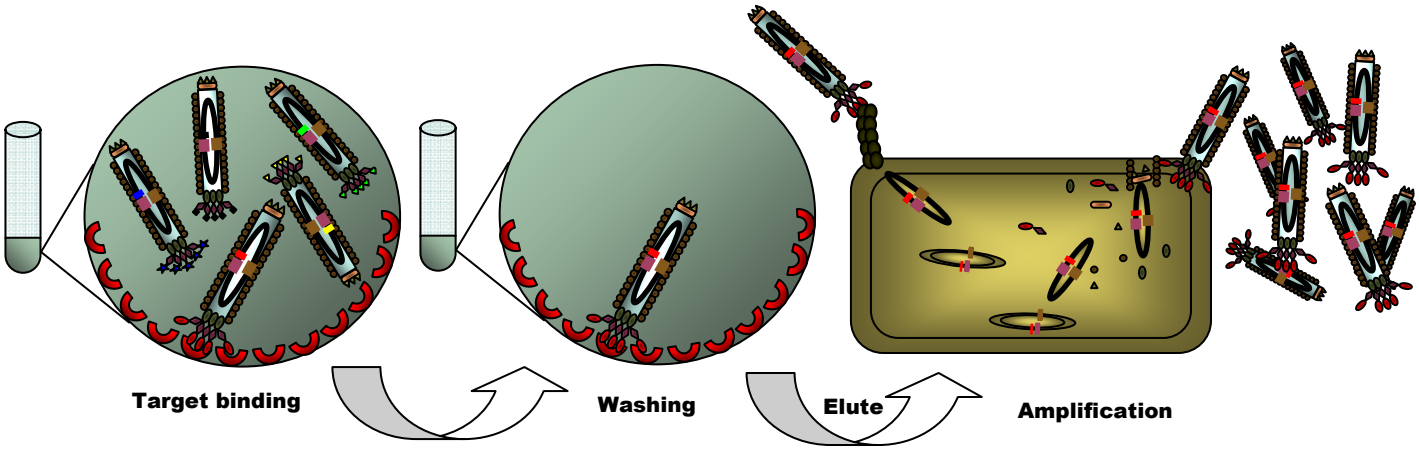


Figure 12. Affinity selection: the general cycle of phage display
- Left panel: A phage library displaying an immense variety of peptides or proteins is exposed to immobilized target molecules of interest (e.g. receptors) (indicated by red aches) and allowed to bind.
- Centre panel: Non-binding phages and weak binders are removed by stringent washing, leaving only phages with appropriate affinity behind. Binders are eluted by conditions that interrupt the phage-target interaction (e.g. pH modulation).
- Right panel: Bacterial host cells are infected with eluted phages which are thereby amplified.
Successive rounds of panning leads to formation of a less and less diverse population of phages and enrichment for high affinity binders.

After repeated rounds of panning, monoclonal phage populations or pools of phage may be selected and analysed. Eventually genes encoding the peptides with desired binding properties are recovered (i.e. by PCR based methods).

The amplified population, comprising millions of copies of each clone, can then be subjected to further rounds of selection to obtain phages displaying peptides with even higher affinity. Additional variation may be introduced to the pool of phages by mutagenesis in between successive rounds of panning.

Finally, selected clones are investigated in further detail and the inserted DNA, encoding the peptide displayed by the phage particle, is recovered. This allows for subsequent sequencing and cloning in to expression systems yielding free peptides or fusion proteins depending on the desired downstream applications. Further more, synthetic versions of the selected peptide may be constructed based on the revealed DNA sequence.

Two key parameters to keep in mind when screening phage display libraries are, stringency and yield. Stringency is the degree to which peptides with higher affinity (or some other preferred ability) are favoured over peptides with lower affinity. Yield is the fraction of peptide expressing phage within a given affinity window that survives the selection procedure. The ultimate goal in affinity selection usually is to retrieve phages that express peptides with high affinity for the target of interest.

Increased stringency (i.e. reduced target concentration, shorter target-library incubation period or excessive washing), may appear to do the trick. However, the immense library diversity and low number of identical clones may cause loss of potential high affinity binders as the number of phage recovered after the first round of panning typically is in the 10^6 vicinity. In the first round of panning good yield of high affinity peptides is of great importance, thus the stringency of the panning procedure must be sacrificed. Prior to subsequent panning rounds, the survivors of the first panning are amplified to millions of identical clones. At this stage high stringency can be applied as yield can safely be decreased. The level of stringency applied is limited however. Panning procedures are imperfect, causing an unavoidable background of phage binding to the solid support (reviewed in Smith and Petrenko, 1997; Willats, 2002).

3. Targeting pIgR: specific delivery to epithelia and mucosa

Oral and systemic delivery of therapeutic macromolecules, gene therapy vectors and diagnostics to tissue may be hampered by endothelial or epithelial barriers. In an effort to overcome this obstacle, delivery strategies exploiting innate trans-barrier transport mechanisms have been suggested. Transport of phage particles across epithelial cells are facilitated by peptides selected from a random phage display peptide library, suggesting that small peptide carriers are capable of binding transcytosing receptors and mediate delivery of macromolecules, gene vectors or liposomes to epithelial cells and surfaces (Ivanenkov and Menon, 2000; Scherrmann, 2002, review).

The pIgR-mediated transcytosis pathway for secretory Abs provides a transport mechanism that can potentially be exploited for the specific delivery of drugs, antibodies, antibody derivatives, viruses, and genes to epithelial cells and mucosal surfaces. pIgR specific scFvs, conjugated to a therapeutically active protease inhibitor, have been shown to be specifically transcytosed across epithelial cells in an *in vitro* model (Eckman et al., 1999; Ferkol et al., 2000). Recently, an *in vivo* epithelial model provided further supportive evidence for the feasibility of utilizing pIgR for specific delivery of therapeutics to epithelial surfaces (Ferkol et al., 2003). Furthermore, pIgR specific Fabs and scFvs have been developed for receptor mediated gene delivery to epithelial cells (Ferkol et al., 1993; Gupta et al., 2001).

An alternative strategy, approached in parallel to scFv and Fab based targeting, is the use of short and simple peptides. Small peptides with high affinity for the polymeric Ig receptor that are transported across epithelial cells by transcytosis may meet the criteria as ideal fusion partners for receptor mediated transfer to epithelia or mucosal surfaces. Selection of viable small peptides candidates, preferably peptides that do not compete with sAbs for pIgR binding, may be achieved through phage display. White and Capra (2002) recently reported selected peptides, both based on the IgA Cα3 sequence and candidates selected from random phage display libraries, that are specifically transcytosed across polarized, pIgR transfected, MDCK cells. Their study supports evidence that small pIgR-binding peptide motifs may deliver therapeutics to mucosal sites.

References

Brandtzaeg P. (1974) Mucosal and glandular distribution of immunoglobulin components: differential localization of free and bound SC in secretory epithelial cells. *J Immunol* **112**, 1553-9.

Brandtzaeg P. (2003) Mucosal immunity: integration between mother and the breast-fed infant. *Vaccine* **21**, 3382-8.

Brock S. C., McGraw P. A., Wright P. F. and Crowe J. E., Jr. (2002) The human polymeric immunoglobulin receptor facilitates invasion of epithelial cells by Streptococcus pneumoniae in a strain-specific and cell type-specific manner. *Infect Immun* **70**, 5091-5.

Braathen R., Sorensen V., Brandtzaeg P., Sandlie I. and Johansen F. E. (2002) The carboxyl-terminal domains of IgA and IgM direct isotype-specific polymerization and interaction with the polymeric immunoglobulin receptor. *J Biol Chem* **277**, 42755-62.

de Oliveira I. R., de Araujo A. N., Bao S. N. and Giugliano L. G. (2001) Binding of lactoferrin and free secretory component to enterotoxigenic Escherichia coli. *FEMS Microbiol Lett* **203**, 29-33.

Eckman E. A., Mallender W. D., Szegletes T., Silski C. L., Schreiber J. R., Davis P. B. and Ferkol T. W. (1999) In vitro transport of active alpha(1)-antitrypsin to the apical surface of epithelia by targeting the polymeric immunoglobulin receptor. *Am J Respir Cell Mol Biol* **21**, 246-52.

Ferkol T., Cohn L. A., Phillips T. E., Smith A. and Davis P. B. (2003) Targeted delivery of antiprotease to the epithelial surface of human tracheal xenografts. *Am J Respir Crit Care Med* **167**, 1374-9.

Ferkol T., Eckman E., Swaidani S., Silski C. and Davis P. (2000) Transport of bifunctional proteins across respiratory epithelial cells via the polymeric immunoglobulin receptor. *Am J Respir Crit Care Med* **161**, 944-51.

Ferkol T., Kaetzel C. S. and Davis P. B. (1993) Gene transfer into respiratory epithelial cells by targeting the polymeric immunoglobulin receptor. *J Clin Invest* **92**, 2394-400.

Gupta S., Eastman J., Silski C., Ferkol T. and Davis P. B. (2001) Single chain Fv: a ligand in receptor-mediated gene delivery. *Gene Ther* **8**, 586-92.

Hammerschmidt S., Talay S. R., Brandtzaeg P. and Chhatwal G. S. (1997) SpsA, a novel pneumococcal surface protein with specific binding to secretory immunoglobulin A and secretory component. *Mol Microbiol* **25**, 1113-24.

Hammerschmidt S., Tillig M. P., Wolff S., Vaerman J. P. and Chhatwal G. S. (2000) Species-specific binding of human secretory component to SpsA protein of Streptococcus pneumoniae via a hexapeptide motif. *Mol Microbiol* **36**, 726-36.

Iannolo G., Minenkova O., Petruzzelli R. and Cesareni G. (1995) Modifying filamentous phage capsid: limits in the size of the major capsid protein. *J Mol Biol* **248**, 835-44.

Ivanenkov V. V. and Menon A. G. (2000) Peptide-mediated transcytosis of phage display vectors in MDCK cells. *Biochem Biophys Res Commun* **276**, 251-7.

Johansen F. E., Braathen R. and Brandtzaeg P. (2000) Role of J chain in secretory immunoglobulin formation. *Scand J Immunol* **52**, 240-8.

Kaetzel C. S. (2001) Polymeric Ig receptor: defender of the fort or Trojan horse? *Curr Biol* **11**, R35-8.

Lowman H. B., Bass S. H., Simpson N. and Wells J. A. (1991) Selecting high-affinity binding proteins by monovalent phage display. *Biochemistry* **30**, 10832-8.

Luton F. and Mostov K. E. (1999) Transduction of basolateral-to-apical signals across epithelial cells: ligand-stimulated transcytosis of the polymeric immunoglobulin receptor requires two signals. *Mol Biol Cell* **10**, 1409-27.

MacDonald T. T. (2003) The mucosal immune system. *Parasite Immunol* **25**, 235-46.

Marvin D. A. (1998) Filamentous phage structure, infection and assembly. *Curr Opin Struct Biol* **8**, 150-8.

Mayer L. (2003) Mucosal immunity. *Pediatrics* **111**, 1595-600.

O'Neil K. T., Hoess R. H., Jackson S. A., Ramachandran N. S., Mousa S. A. and DeGrado W. F. (1992) Identification of novel peptide antagonists for GPIIb/IIIa from a conformationally constrained phage peptide library. *Proteins* **14**, 509-15.

Phalipon A. and Corthesy B. (2003) Novel functions of the polymeric Ig receptor: well beyond transport of immunoglobulins. *Trends Immunol* **24**, 55-8.

Riechmann L. and Holliger P. (1997) The C-terminal domain of TolA is the coreceptor for filamentous phage infection of E. coli. *Cell* **90**, 351-60.

Scherrmann J. M. (2002) Drug delivery to brain via the blood-brain barrier. *Vascul Pharmacol* **38**, 349-54.

Scott J. K. and Smith G. P. (1990) Searching for peptide ligands with an epitope library. *Science* **249**, 386-90.

Sidhu S. S. (2001) Engineering M13 for phage display. *Biomol Eng* **18**, 57-63.

Smith G. P. (1985) Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface. *Science* **228**, 1315-7.

Smith G. P. and Petrenko V. A. (1997) Phage Display. *Chem Rev* **97**, 391-410.

Spiekermann G. M., Finn P. W., Ward E. S., Dumont J., Dickinson B. L., Blumberg R. S. and Lencer W. I. (2002) Receptor-mediated immunoglobulin G transport across mucosal barriers in adult life: functional expression of FcRn in the mammalian lung. *J Exp Med* **196**, 303-10.

White K. D. and Capra J. D. (2002) Targeting mucosal sites by polymeric immunoglobulin receptor-directed peptides. *J Exp Med* **196**, 551-5.

Willats W. G. (2002) Phage display: practicalities and prospects. *Plant Mol Biol* **50**, 837-54.

Zacher A. N., 3rd, Stock C. A., Golden J. W., 2nd and Smith G. P. (1980) A new filamentous phage cloning vector: fd-tet. *Gene* **9**, 127-40.

Zhang J. R., Mostov K. E., Lamm M. E., Nanno M., Shimida S., Ohwaki M. and Tuomanen E. (2000) The polymeric immunoglobulin receptor translocates pneumococci across human nasopharyngeal epithelial cells. *Cell* **102**, 827-37.

Manuscript

Identification and characterization of phage displayed, SC binding peptides, with potential for use in pIgR mediated mucosal targeting

Anders Sandvik [1], Ranveig Braathen [2], Gøril Berntzen [1], Finn-Eirik Johansen [2], Inger Sandlie [1] and Vigdis Lauvrak [1*]

[1] University of Oslo, Department of Molecular Bioscience, Division of Cell and Molecular Biology, P.O.box 1050, Blindern, 0316 Oslo, Norway

[2] University of Oslo, Institute of Pathology, Laboratory for Immunohistochemistry and Immunopathology, Rikshospitalet, 0027 Oslo, Norway

*To whom correspondence should be addressed

Key words: Mucosal targeting, pIgR, Phage display, Automation, Amino acid substitution, Mutation analysis, GFP-fusion protein.

Abbreviations

A, Adenine

A_{405}, Absorption at 405nm

ABTS, 2,2'-Azino-bis (3-ethylbenzthia-zoline-6-sulfonic-acid)

AMP, Adenosine monophosphate

AP, Alkaline phosphatase

Bio, Biotin

BSA, Bovine serum albumin

C, Cytosine

C1q, Complement system factor 1 subcomponent q

C-terminal, COOH-terminal

dH_2O, Distilled water

dNTP, deoxy-N-triphosphate (where N=adenine, guanine, thymine or cytosine)

E, Eluate

EA, Amplified eluate

ELISA, Enzyme linked immunosorbent assay

fSC, Free SC

G, Guanine

GFP, Green fluorescent protein

h(rs), Hour(s)

HRP, Horse radish peroxidase

Ig, Immunogloublin

IPTG, Isopropyl-β-D-thiogalactoside

kDa, Kilo Daltons

LB, Luria Bertani medium

LB/amp, LB with 50µg/mL ampicilline

LB/tetc, LB with 20µg/mL tetracycline

mAb(s), Monoclonal antibody(s)

MDCK, Madin-Darby canine kidney

min, Minute(s)

NEB, New England BioLabs

N-terminal, NH2-terminal

ON, Over night

PBS, Phosphate buffer saline

PBS/T, PBS with 0,05% Tween 20

PCR, Polymerase chain reaction

pIgA, Polymeric immunoglobulin A

pIgM, Polymeric immunoglobulin M

pIgR, Polymeric Immunoglobulin receptor

RF, Replicative form

RFU, Relative fluorescence units

rpm, Rotations per minute

RT, Room temperature

SB, Super Broth medium

SC, Secretory component

scFv, Single chain fragment of variation

SDS PAGE, Sodium dodecyl sulphate polyacrylamide gel electrophoresis

sec, Seconds

sIgA, Secretory IgA

sIgM, Secretory IgM

SKM, Skimmed milk

T, Thymine

TAE, Tris-acetat EDTA

TU, Transducing units

U, Units

wt, Wild type

α, Anti

1. Abstract

Delivery of therapeutic macromolecules, gene therapy vectors and diagnostics to tissue may be hampered by endothelial or epithelial barriers. In an effort to overcome this obstacle, delivery strategies exploiting innate trans-barrier transport mechanisms have been suggested. The polymeric immunoglobulin receptor (pIgR) directs formation and trans-epithelial transportation of secretory immunoglobulins that provides a first line of humoral defence at mucosal surfaces. Consequently, the pIgR-mediated mucosal secretion system potentially represents a route for targeted delivery to epithelial cells and mucosal surfaces. We aimed at identifying novel small peptides that bind to the extracellular domains of pIgR, known as the secretory component (SC), and that are specifically transcytosed across Madin Darby Canine Kidney (MDCK) cells transfected with human pIgR, by affinity selection from phage display libraries. Assuming that one key to identifying peptides, with the sought after properties, is the number of screened phage single clones, semi-automated screening procedures were designed. We have identified 5 novel SC binding phage displayed peptides by affinity selection using two libraries. One of the identified clones, C9-4, bind immobilized SC in a SpsA-like manner and is specifically transcytosed across MDCK cells transfected with human pIgR. Amino acid substitution mutation analysis revealed that all, but three, amino acids were indispensable to the selected peptide's functionality as an SC binder. Furthermore we constructed and expressed C9-4 as a fusion to green fluorescent protein (GFP). Despite the successful affinity selection, transcytosis, and fusion protein construction and expression, interaction with immobilized SC could not be demonstrated in the new fusion format.

2. Introduction

Mucosal surfaces, constantly challenged with potential pathogens, are protected by innate and adaptive immune defense mechanisms. Adaptive immunity at the mucosa is mediated mainly by the polymeric immunoglobulins (pIgs). Polymeric IgA and IgM contain a small polypeptide called the "joining" (J) chain. J chain-positive IgA primarily consists of dimers (dIgA), and some larger polymers, whereas only IgM pentamers (pIgM) incorporate the J chain (reviewed in Johansen et al., 2000). dIgA and pIgM bind to the polymeric immunoglobulin receptor (pIgR) expressed on the basolateral side of mucosal epithelial cells. The protective antibodies are transported across the epithelia by pIgR to mucosal surfaces (Norderhaug et al., 1999, review). pIgR is a transmembrane type I receptor consisting of five Ig homologous domains a membrane spanning domain and a cytoplasmic tail. Internalization signals in the cytoplasmic, C-terminal domain of pIgR causes the receptor to be sorted into coated pits and hence endocytosed. In the endosomal compartment, pIgR is sorted into vesicles for transcytosis to the apical surface or recycled to the basolateral surface. At the apical surface, a protease cleaves off the extracellular receptor domain, named secretory component (SC). pIgR is constitutively transcytosed, regardless of ligand binding. However, ligand binding increases the internalization rate (Luton et al., 1999). Consequently, free secretory component (fSC) as well as SC associated pIgs is released into the extracellular secretions. SC associated Abs are referred to as secretory Abs (sAb) i.e. secretory IgA (sIgA) and IgM (sIgM).

sIgA, by far the dominating species of sAbs, remains covalently linked to SC in secretions. Pentameric IgM, on the other hand, remains non-covalently associated with SC in secretions after apical cleavage. Anchored to the mucus layer by SC, secretory antibodies serve as an external barrier that performs immune exclusion at mucosal surfaces, thereby protecting the epithelium from pathogen colonization and invasion. Furthermore, dIgA bound to pathogens located in the lamina propria are transported through the epithelia by pIgR and are excreted at the luminal face as complexes of sIgA and antigen. En route to luminal secretion, pIgA bound to pIgR may counteract invading pathogens by intracellular neutralization and subsequent excretion (reviewed in Phalipon and Corthesy, 2003). Uncleaved pIgR is internalized and recycled apically (Breitfeld et al., 1989). pIgR, recycling at the apical face of the epithelia, is suggested to provide a mechanism for colonization and invasion of the epithelia.

SpsA (also known as CbpA) is a surface displayed protein virulence factor of *Streptococcus pneumoniae*. *S. pneumoniae* is a potentially lethal bacterial pathogen that commonly colonizes the upper respiratory epithelium in an asymptomatic fashion. To invade, the pathogen must cross the respiratory epithelial barrier, and it has been shown to bind human pIgR, fSC and sIgA (Hammerschmidt et al., 2000; Zhang et al., 2000). Expression of pIgR in epithelial model systems has been shown to greatly enhance pneumococcal adherence and invasion, suggesting that pIgR recycling is one of several mechanisms utilized for invasion (Zhang et al., 2000).

In light of the pIgR biology, the pIgR-mediated transcytosis pathway is a mechanism that might be exploited for the specific delivery of therapeutic payloads, diagnostic agents or gene delivery vectors to epithelial cells and mucosal surfaces. Furthermore, pIgR binding peptides might have the potential to block ligand binding sites on the receptor, potentially leading to reduced transcytosis or entry of pathogens.

Phage display libraries have successfully been the source of peptides with affinity for a large variety of targets. Peptides used in detection of biological threats (Petrenko and Vodyanoy, 2003, review), mapping of molecular markers of tumors (Zurita et al., 2003, review), mapping of enzyme specificity (Deperthes, 2002, review), vaccine development (Lesinski and Westerink, 2001, review) and targeted drug delivery, i.e. to tumors (Nilsson et al., 2000, review), have been performed using phage display libraries. Functional mimicry of a protein hormone, by a peptide agonist that bind the EPO receptor (Livnah et al., 1996), demonstrates that small peptides without resemblance to the receptor's native ligand, may be biologically active. Phage displayed peptides have been demonstrated to block receptors. Hetian et al., (2002) describe a peptide that block the vascular endothelial growth factor receptor and thus have potential in inhibition of angiogenesis and tumor growth. Mucosae have been targeted both by integrin binding phage displayed peptides (Ivanenkov and Menon, 2000) and through pIgR binding peptides. Recently, pIgR targeting peptides were identified from phage display libraries and demonstrated to be specifically transcytosed across an epithelial model system (White and Capra, 2002).

Furthermore, four peptides that bind immobilized SC were isolated by affinity selection from two cysteine constrained phage display libraries (Lauvrak et al., 2003; Braathen et al., unpublished data).

Only one of the identified peptides, named C9-4, was specifically transported across MDCK cells transfected with human pIgR. The peptide was shown to bind to SC at a site different from the interaction site utilized by IgA and IgM. Interestingly, phages displaying the peptide were inhibited from target binding by SpsA. Inhibition studies using a synthetic version of the C9-4 peptide interfered with the phage-receptor interaction at nM concentrations (Braathen et al., unpublished results).

Assuming that one key to isolating more peptide candidates was to screen additional peptide displaying phage clones, semi-automated screening procedures to allow for high throughput sample handling and screening was set up. Access to a robotic liquid handler and large amounts of human SC purified from colostrum, made SC a feasible target to use when designing an automation pilot project. Furthermore, to study the interaction between the C9-4 peptide and the receptor, the peptide sequence was analyzed by systematically substituting each amino acid.

Phage displayed peptides affinity selected for various targets have been attempted fused to scFv- and Fab antibody fragments in our laboratory. One of these peptides, a C1q binder, retained its target binding ability both as a C-terminal fusion to Fab, scFv and as an internal fusion to Fab (Lauvrak and Berntzen, unpublished). A scFv antibody fragment with C9-4 as C-terminal fusion and a green fluorescent protein (GFP) with C9-4 as N-terminal fusion have previously been engineered and expressed in bacteria. However, both constructs gave low yields and no SC binding was observed using ELISA assays (Lauvrak and Berntzen, unpublished). Lack of binding in the new fusion protein context can be due to low accessibility, low affinity and/or improper folding of the peptide. To investigate if the introduction of a flexible linker and dimerization of the C9-4 peptide would affect pIgR binding, through increased peptide accessibility and avidity, N-terminal GFP fusion proteins were constructed.

The goal of the current study has been to identify and isolate small peptide motifs that would both bind to the pIgR and be specifically transcytosed by the receptor and thus offer the possibility of targeting mucosal sites. We aimed at developing efficient screening methods for handling of large numbers of phage single clones. Moreover, we have studied the C9-4 peptide's interaction with pIgR in further detail.

3. Materials and Methods

3.1 Phage display libraries, phage clones, plasmids and bacterial strains

The fUSE5 vector (Scott and Smith, 1990) was kindly provided by Dr. G.P. Smith, University of Missouri, USA. The phage protein III fusion libraries Cys9 (Lauvrak et al., 2003), and 7-mer (Lauvrak, unpublished) were both constructed in our lab using the fUSE5 vector. The Cys9 library consisted of approximately 5 x 10^7 different clones and displays constrained peptides of nine random amino acids, flanked by cysteines. The 7-mer library consisted of approximately 8 x 10^7 different clones and displays random unconstrained peptides, seven amino acids in length. The libraries were constructed from degenerative synthetic oligonucleotides (NNK), where N represents an equal mix of the four natural DNA nucleotides; Adenine (A), guanine (G), cytosine (C) and thymine (T) and K represents an equal mix of G and T. Phage clone C9-4 (CVVWMGFQQVC) was previously selected from the Cys9 library and shown to bind SC and demonstrated to be transcytosable (unpublished results, Braathen et al.). The irrelevant phage clones T7 (LHVSQRG) and TCys9 (CDLGWWMDVSC) were randomly chosen from the 7-mer and Cys9 libraries, respectively.

The bacterial expression vector pET21d was obtained from Novagen. The green fluorescent protein (GFP) encoding vector pHOG-GFP was kindly provided by Dr. O.H. Brekke, Affitech AS, Norway, and is a modified version of the bacterial expression vector pHOG (Kipriyanov et al., 1997) with a GFP encoding sequence inserted C-terminally of an scFv encoding sequence (Brekke, unpublished). The bacterial strains *E.coli* BL21DE and *E.coli* XL-1 blue were obtained from Novagen and Stratagene, respectively. *E.coli* K91K (Smith and Scott, 1993) and *E.coli* MC1061 (Meissner et al., 1987) were both kind gifts from Dr. G.P. Smith, University of Missouri, USA.

3.2 Preparation and transformation of electrocompetent E.coli cells

10mL Super broth growth medium (SB) was inoculated with *E.coli* MC1061, *E.coli* XL-1 blue or *E.coli* BL21DE cells and incubated at 37^{o}C, over night (ON), with agitation (220rpm) (*Innova*^{TM} 4000 incubator shaker, New Brunswick Scientific).

One litre of preheated SB with 5mL of glucose (2M), in a three litre shaking flask, was inoculated with 5mL of the starter culture. The *E.coli* MC1061 cultures were grown in the presence of 100µg/mL streptomycin and *E.coli* XL-1 blue cultures were prepared using 10µg/mL tetracycline. The cultures were incubated at 37°C with agitation (220rpm) (Controlled environment incubator shaker, New Brunswick Scientific) until a density of 0.8 at OD_{600} was reached (PU 8750 UV/VIS scanning spectrophotometer). Shaking flasks were cooled on ice for 15 minutes (min). Next, cells were harvested by centrifugation at 5000rpm for 15min at 4°C (*Sorvall*® RC-5B Refrigerated superspeed centrifuge, DuPont Instruments). The cell pellet was washed 3x by gentle re-suspension in ice cold dH_2O with 10% glycerol. The amount of washing reagent was reduced between each round of washing (500mL, 250mL and 50mL). Each washing step was concluded with centrifugation as described above. After the last washing step, the pellet was gently resuspended in 2mL of ice cold dH_2O with 10% glycerol. Aliquots of 40µL in cold tubes were immediately frozen in ethanol- dry ice bath and stored at -70°C.

Electrocompetent bacteria were transformed by electroporation. 2µL of the vector DNA to be propagated or 2µL of *seeDNA*^{TM} (Amersham Bioscience) precipitated ligation mix, were mixed with 40µL electrocompetent cells on ice and transferred to ice cold cuvettes (1mm gap, BTX). Subsequently, cells were pulsed (BioRad, 200Ω, 1.3V), recovered from the cuvette by adding 1mL 2xYT growth medium, transferred to an eppendorf tube and incubated at 37°C for 1hour (h). *E.coli* MC1061 transformed with fUSE5 was recovered and incubated using 2xYT growth medium containing tetracycline (0.2µL/mL). The low concentration of antibiotic, does not affect sensitive cells, however it induce expression of the tetracycline resistance gene. Cell dilutions were then plated on to LB plates with relevant antibiotic and incubated at 37°C ON. *E.coli* MC1061 transformed with fUSE5 were plated onto tetracycline (20µg/mL) containing plates (LB/tetc), whereas *E.coli* XL-1 blue transformed with pHOG-GFP and *E.coli* BL21DE transformed with pET21d were plated onto LB plates containing ampiciline (50µg/mL) (LB/amp).

3.3 Phage titration, amplification and propagation of single clones

The number of phage particles in a solution was determined by titration as *E.coli* K91K transducing units per millilitre (TU/mL), i.e. the number of phage particles that infect host cells, calculated as tetracycline resistant colonies.

fUSE5 phages have a reduced infectivity rate (the ratio of TU to the number of physical phage particles) of ~5-10%, compared to ~50% infectivity in wild type phage (Smith, 1994). Thus, the actual amount of phage particles in a sample may exceed the TU/mL value by a factor of 10-20.

Phage amplification was done by growing transduced *E.coli* K91K cells in liquid media and phage containing supernatants were obtained by removing the cells by centrifugation. Due to a defect in minus strand synthesis, and consequently low copy number of RF-DNA, fUSE5 can be propagated for an extended period of time without causing detrimental harm to the host cells (Smith, 1994).

Fresh Terrific Broth (TB) cultures of *E.coli* K91K cells were prepared prior to each titration or amplification by inoculating 9mL TB medium added 1mL phosphate buffer (0.17M KH_2PO_4 and 0.72M K_2HPO_4) with a single colony of *E.coli* K91K. Cells were incubated for 4 hours (hrs) at 37^{o}C with agitation (200rpm) (*Innova*^{TM} 4000 incubator shaker, New Brunswick Scientific) to yield a high density of living host cells (~10^9cells/mL). The last 15 minutes of the incubation time, the shaking rate was reduced to 50 rpm to allow for regeneration of F-pili.

Dilutions of the phage suspensions to be titrated, was prepared in PBS. 10µL of diluted phage was mixed with 10µL of *E.coli* K91K TB culture prepared and incubated for 10min at RT. The ratio of phage TU versus host cells was always assured to be below 0.1. Luria Bertani medium (LB) with tetracycline (0.2µg/mL) was added to a total of 200µL and the mixture was incubated for 45min at 37^{o}C to allow for gene expression from the tetracycline resistance gene. Subsequently, 100µL of the cultures were plated onto LB/tetc plates and incubated at 37^{o}C ON. Colonies were counted to calculate TU/mL.

Phages were amplified by mixing 100µL phage supernatant with 100µL of fresh *E.coli* K91K TB host cells and incubated as for the titration, using 1mL of LB medium with low tetracycline concentration. Subsequently, the cultures were transferred to 20mL of tetracycline (20µg/mL) added LB medium (LB/tetc) in a 50mL tube. Dilutions of the 20mL culture were plated onto LB/tetc plates and incubated at 37^{o}C ON to determine output. The 20mL culture was incubated at 37^{o}C ON with agitation. Cells were removed by centrifugation at 4600rpm for 30min at 4^{o}C (Centra CL3, International Equipment Company).

The phage supernatant was centrifuged once more to remove remaining cells and cell debris. Phage supernatants were stored at 4°C.

Individual phage clones were propagated and screened for target binding capacity. Fresh colonies of transduced *E.coli* K91K cells, prepared as described above, were transferred to 15mL tubes containing 3mL LB/tetc or to 96 deep well plates containing 900µL of LB/tetc by standard sterile techniques and incubated at 37°C, ON with agitation. Phage supernatants were recovered by two centrifugations as described above, or for 30min at 4000rpm at 4°C (*Allegra*^{TM}25R Centrifuge, Beckman Coulter) for the 96 deep well cultures. The phage supernatants were transferred to 15mL tubes or 96well V-bottom plates and stored at 4°C.

3.4 Affinity selection for SC binding, peptide displaying, phages

Novel SC binding peptides were isolated by performing two rounds of affinity selection using the 7-mer and Cys9 libraries. SC purified from human colostrum was provided by Dr. F-E. Johansen, Laboratory for Immunohistochemistry and Immunopathology, Institute of Pathology, University of Oslo, Rikshospitalet. *Nunc-Immuno*^{TM} *MaxiSorp*^{TM} tubes, with "star" shaped bottom, or *Nunc-Immuno*^{TM} *MaxiSorp*^{TM} strips were coated with 800µL or 200µL of 30µg/mL human SC diluted in PBS, respectively by ON incubation at 4°C. The tubes and strips were blocked using two coat volumes of either 1% skimmed milk (SKM) or 1% bovine serum albumin (BSA) (Sigma-Aldrich) in PBS at room temperature (RT) for 1h followed by three times washing in PBS with 0.05% Tween20 (Sigma-Aldrich) (PBS/T). Input phage to the first round of affinity selection was 100µL of the Cys9 library (~$1.4x10^{10}$TU) and 10µL of the 7-mer library (~$1.5x10^{9}$TU), diluted to 800µL or 200µL, respectively, in 1% SKM or 1% BSA and pre-incubated for 30min at RT before they were added to the coated and blocked tubes or wells. Phages were allowed to interact with and bind to immobilized SC for 90min at RT. Following washing, 8x PBS/T, bound phages were eluted by adding 800µL or 200µL 0.1M HCl/Glycine pH 2.2 to the tubes and wells of the strips, respectively. The acidic buffer was incubated for 10min at RT before the phage containing eluates were neutralized by adding 120µL or 17µL of 1.5M Tris pH8.8 buffer, respectively. 200µL of the first eluates (E1) were amplified in 20mL cultures by infecting fresh *E.coli* K91K as described in section 3.3. The output of phage from each selection was determined as the total number of TUs in the eluates. 100µL portions of the amplified eluates (EA1) served as input to the second round of affinity selection.

Phage supernatant nomenclature:
Eluates from the first and second round of selection are referred to as E1 and E2, respectively. Amplified eluates from the first and second round of selection are referred to as EA1 and EA2, respectively. The prefix 7-mer or Cys9 refers to the library of origin. The applied blocking agent is included in the eluate/amplified eluate name via the suffix BSA or SKM.

3.5 Phage ELISA

Supernatants, originating from pools of phage or single clones, were analyzed for the presence of SC binding particles by ELISA assays with immobilized SC as target. *Nunc*^{TM} *MaxiSorp*^{TM} 96 well plates were coated with 100µL or 200µL of 5µg/mL SC by incubation at 4^{o}C, ON. Wells were blocked, using two coat volumes of 1% SKM or 1% BSA, for 1h at RT, and subsequently washed 3x in PBS/T. Phage supernatants, 50µL or 100µL, were pre-incubated in blocking solution (1:1) for 30min at RT. Pre-incubated phages (100µL or 200µL) were allowed to interact with, and bind to immobilized SC for 90min at RT. Following extensive washing, 8x PBS/T, bound phages were incubated with one coat volume of M13 specific mAbs (from mouse) conjugated to horse raddish peroxidase (αM13-HRP) (Amersham Pharmacia Biotech, 5000x), diluted (1:4000) in blocking solution, for 1h at RT. The washing step was repeated before bound phages were visualized. One coat volume of ABTS (Sigma-Aldrich), dissolved in sterile filtrated 0,05M citric acid buffer and added H_2O_2 (Aldrich, 30%), was applied as HRP substrate. Readout was absorbance at 405nm (A_{405}) as measured (Tecan Sunrise Spectrophotometer, with Magellan 3.0 software) after 30-60 minutes of incubation at RT.

3.6 Phage elution-titration study

The number of phages bound to SC was determined by elution-titration studies carried out in parallel to ELISA analysis. *Nunc*^{TM} *MaxiSorp*^{TM} 96 well plates were coated with 200µL of 20µg/mL SC by incubation at 4^{o}C, ON. Subsequently, wells were blocked and washed and phages were incubated as described for the ELISA assay in 3.5. Washing, elution with 200µL of 0.1M HCl/Glycine pH 2.2 and neutralization using 30µL of 1.5M Tris pH8,8 buffer were performed essentially as described for affinity selection in 3.4. Output of recovered phages was determined by infecting fresh *E.coli* K91K cells, as described in 3.3., to validate enrichment for SC binding clones.

3.7 Phage sequencing

gpIII displayed fusion peptides, of ELISA positive phages, were characterized by recovering the DNA sequences by PCR. A large number of sequencing templates were prepared in a 96 well format, using a Biomek 2000 liquid handling robot.

Each PCR reaction contained: 10µL phage supernatant as template, primers fUSE5 (TCGAAAGCAAGCTGATAAACCG) and fUSE5-for (GTACAAACCACAACGCCTGTAG) (20pmol) (Eurogenetic), dNTP (200µM, Amersham Pharmacia Biotech), 10µL Thermopol polymerase buffer (10x, NEB), 0.5µL Vent DNA polymerase (2000U/mL, New England BioLabs (NEB)) and dH_2O to a total of 90µL. Templates and PCR reaction mix were transferred to 96well V-bottom PCR plates using an automated liquid handler, Biomek 2000, (Beckman Coulter) (see appendix 1 for robotic scripts) or manually. The sequencing templates were amplified by the following cycles: 94^{o}C for 5min, then 30 cycles of 1min 94^{o}C, 55^{o}C for 30 seconds (sec) and 72^{o}C for 30sec using a Techne Genius thermocycler (Techne). PCR products were purified (*QIAquick*® PCR Purification Kit, Qiagen) as described in the protocol provided with the kit. Purified PCR products were analysed by gel electrophoresis (BioRad powerpac 300 and Pharmacia electrophoresis power supply 500/400, 90-100V) (2% agarose gel, Sigma-Aldrich, in TAE buffer), to confirm generation of sequencing templates. Agarose gels were prepared with casting combs designed to accommodate sample application using multi-channel pipettes, consequently leading to enhanced throughput.

Purified templates were shipped to GATC Biotech AG for sequencing using primer sekP (CCCTCATAGTTAGCGTAACG), complementary to the fUSE5 + strand. The sequencing results were provided electronically as abi files and analysed by Chromas software (McCarthy).

3.8 Inhibition of phage binding to SC

Phage binding to SC was further examined by the addition of competing target binding molecules in an SC binding ELISA. The native ligands of human SC, pIgA and pIgM plus fragments of pneumococcal surface protein SpsA (*Streptococcus pneumoniae*) that bind human SC species specifically were applied as inhibitors. pIgA and pIgM was prepared and purified as described by Braathen et al., (2002). SpsA fragments SH2, SM1 (SC positive) and SH3 (SC negative) (Hammerschmidt et al., 1997; Hammerschmidt et al., 2000) were obtained from S. Hammerschmidt. *Nunc*^{TM} *MaxiSorp*^{TM} 96 well plates were coated with 50µL of 30µg/mL SC by incubation at 4^{o}C, ON. Subsequently, wells were blocked, using two coat volumes of 1% BSA, for 1h at RT and washed in PBS/T (4x using a Tecan Columbus washer, Tecan). 60µL of inhibitor dilutions (end concentration 100, 50, 25, 12.5 and 0nM) were added to 60µL phage supernatant. 50µL samples were incubated against immobilized SC of for 90min at RT. The washing step was repeated prior to incubation with one coat volume of mouse anti-M13 antibody (Amersham Pharmacia Biotech) diluted 1:1000 in PBS/T for 90 minutes at RT. Following repeated washing, 50µL of the secondary antibody reagent, rabbit anti-mouse conjugated to alkaline phosphatase (AP) (DAKO) diluted 1:1000 in PBS/T was added for 90 min at RT. Finally, following washing, the plates were incubated with *p*-nitrophenyl phosphate (Sigma) in diethanolamine buffer at RT for 10-60 min. Absorbance at 405 nm was measured as described in 3.5.

3.9 Transcytosis of phage across pIgR transfected MDCK cells

Madin-Darby canine kidney (MDCK) cells, a commonly used epithelia model system, were used to investigate if SC binding phage, identified by affinity selection, could be transported across epithelia by pIgR mediated transcytosis.

MDCK cells, untransfected or stably transfected with human or murine pIgR (Natvig et al., 1997; Roe et al., 1999) were used to study transcytosis. Approximately 5.0×10^5 cells were seeded on 1-cm^2, 3.0-µm collagen-coated PTFE filters (Transwell-COL 3494; Costar) The cells were incubated for 3 days at 37 °C with 5% CO_2 in Dulbecco's modified Eagle's medium containing 10% fetal calf serum, 50 µg/ml gentamicin, and 1 mM L-glutamine. At that time, the transepithelial resistance was about 150-200 ohms.

10 µl of phage supernatant were added to the basolateral chamber and the filters were incubated for 20 h at 37°C in fresh medium. Irrelevant phage, was used as a control together with 100 nM (15µg/mL) human IgG (to control for unspecific transport and leakage). The apical medium was harvested, and the number of phage particles, transported by pIgR, was calculated by standard titration studies as described in section 3.3. IgG leakage to the apical medium was detected by ELISA essentially as described in Braathen et al., (2002).

3.10 Amino acid substitution mutation analysis

The contribution of each individual amino acid side chain on C9-4's (CVVWMGFQQVC) ability to bind SC was determined by performing an alanine / serine scan on the phage displayed peptide.

3.10.1 Production of inserts encoding amino acid substituted peptides

Synthetic oligonucleotides (Eurogenetic) comprising systematic substitutions of wild type amino acids with alanine or serine were designed with *Sfi*I (highlighted in yellow) restriction sites. The sequences of the mutagenic oligonucleotides are given in table 1. All templates differed only in one codon, except the double mutant AS-C1SC11S that differed from the wild type in two codons.

*sfi*I *sfi*I

5'- CTA TTC TCA CTC GGC CGA CGG GGC C- Peptide insert -GG GGC CGC TGG GGC CGA AAC TGT TGA A- 3'

Table 1. Synthetic oligonucleotides encoding alanine (GCA) or serine (AGC) substitutions

Oligo name	Oligonucleotide sequence
C9-4 *	TGT GTG GTT TGG ATG GGT TTT CAG CAG GTG TGT
AS-C1S	AGC GTG GTT TGG ATG GGT TTT CAG CAG GTG TGT
AS-V2A	TGT GCA GTT TGG ATG GGT TTT CAG CAG GTG TGT
AS-V3A	TGT GTG GCA TGG ATG GGT TTT CAG CAG GTG TGT
AS-W4A	TGT GTG GTT GCA ATG GGT TTT CAG CAG GTG TGT
AS-M5A	TGT GTG GTT TGG GCA GGT TTT CAG CAG GTG TGT
AS-G6A	TGT GTG GTT TGG ATG GCA TTT CAG CAG GTG TGT
AS-F7A	TGT GTG GTT TGG ATG GGT GCA CAG CAG GTG TGT
AS-Q8A	TGT GTG GTT TGG ATG GGT TTT GCA CAG GTG TGT
AS-Q9A	TGT GTG GTT TGG ATG GGT TTT CAG GCA GTG TGT
AS-V10A	TGT GTG GTT TGG ATG GGT TTT CAG CAG GCA TGT
AS-C11S	TGT GTG GTT TGG ATG GGT TTT CAG CAG GTG AGC
AS-C1SC11S	AGC GTG GTT TGG ATG GGT TTT CAG CAG GTG AGC

*C9-4 is included in the table for comparison only.

This template was recovered from a SC binding, phage displayed peptide.

Peptide encoding inserts were produced by PCR amplification of the synthetic oligonucleotides. Each reaction contained: Oligonucleotide template (40pmol), 5'-biotinylated primers P1-bio (CTATTCTCACTCGGCCGACG) and P2-bio (TTCAACAGTTTCGGCCCCAG) (200pmol) (Eurogenetic), dNTP (200µM) and dH_2O to a total of 100µL. Inserts coding for mutations AS-C1S, AS-V2A and AS-W4A were obtained using 0.5µLVent polymerase and 10µL Thermopol polymerase buffer (10x), whereas inserts encoding the remaining substitutions, were acquired using 0.5µL *Herculase*^{TM} Enriched DNA polymerase (5000U/µL)(Stratagene) and 10µL *Herculase*^{TM} reaction buffer (10x) (Stratagene).

The reaction was run under the following conditions: 94°C for 5min, then 20 cycles of 94°C for 1min, 50°C for 30 seconds (sec) and 72°C for 30sec on a Techne Genius thermocycler. The PCR products were purified with *Centriflex*^{TM} protein removal cartridges (Advanced Genetic Technologies Corp.), *seeDNA*^{TM} precipitated as described in the protocol provided with the kit using a microcentrifuge (Heraeus, Biofuge pico), and resuspended in 10µL dH_2O.

Samples of the *seeDNA*^{TM} precipitated PCR products (2µL) were analysed by gel electrophoresis using 4% *NuSieve*®3:1 agarose (BMA products) in TAE buffer.

The double stranded products were restricted with *Sfi*I to generate fUSE5 compatible inserts. Each digestion reaction contained: *seeDNA*^{TM} precipitated PCR product aliquot, 10µL NEB2 buffer (10x), 5µL *Sfi*I (20000U/mL, NEB) and dH_2O to a total of 100µL. The restriction reactions were carried out at 50°C for 4h using a Techne Genius thermocycler. Subsequently, 10µL of the restriction reactions were analysed by gel electrophoresis (4% *NuSieve*®3:1 agarose).

Biotinylated ends and undigested PCR products were removed by incubating two volumes of restriction mix with one volume of *Streptavidin* (Promega, 1mg/mL) for 30min at RT and subsequent capture by *Centriflex*^{TM} protein removal cartridges (Ihle and Michaelsen, 2000). fUSE5 compatible, peptide encoding inserts were recovered using a micro centrifuge (Heraeus, Biofuge pico), *seeDNA*^{TM} precipitated, resuspended in 20µL dH_2O and analyzed on a 4% *NuSieve*®3:1 agarose gel by electrophoresis to determine DNA concentration.

3.10.2 Generation of phage expressing amino acid substituted peptides

fUSE5 RF DNA was propagated by transforming electrocompetent *E.coli* MC1061 cells (competency: ~2x10^{10}Trf/µg pUC DNA) with 2µL RF DNA (~300ng/µL) (diluted 1:1000 in dH_2O) as described in section 3.2. Single clones of the transformants were transferred to 10mL LB/tetc growth medium in 50mL tubes and incubated at 37°C ON with agitation. The following day RF-DNA was isolated from the ON cultures by *Wizard® Plus SV Miniprep* (Promega) as described in the protocol provided in the DNA isolation kit.

fUSE5 RF-DNA was restricted with *Sfi*I to release a 15bp stuffer fragment. Each restriction reaction contained: RF-DNA aliquot (90µL), 20µL NEB2 buffer (10x), 2µL BSA (100x), 2,5µL *Sfi*I and dH_2O to a total of 200µL. The restriction reactions were carried out at 50°C for 4h using a thermocycler (Techne Genius). 5µL additional *Sfi*I diluted in NEB2 buffer (1:20) was added to each reaction tube and the incubation was continued for one hour at the same temperature conditions. Restricted and uncut RF-DNA were quality checked by gel electrophoresis (0.8% agarose), *seeDNA*^{TM} precipitated and concentration evaluated by gel electrophoresis.

Peptide encoding inserts were ligated to restriction digested fUSE5 RF-DNA in a 1:1 molar ratio. Approximately 100ng of fUSE5 vector DNA and 1ng of the various inserts, as determined by gel electrophoresis band intensity, were added to a ligation mix containing 0.5µL T4 DNA ligase (1000U/mL) (Roche), 2µL ligase reaction buffer(10x)(NEB) and dH_2O to a total of 20µL. The ligation reactions were carried out at RT ON, *seeDNA*^{TM} precipitated and resuspended in 5µL dH_2O. All DNA was stored at -20^{o}C.

Phage particles expressing alanine or serine substituted gpIII fusion peptides were generated by transforming *E.coli* MC1061 cells using 2µL of each *seeDNA*^{TM} precipitated ligation mix as described in section 3.2. The transformation cultures (1mL) were transferred to 50mL tubes containing 25mL LB/tetc medium and incubated at 37^{o}C ON with agitation. The following day phage containing supernatants were recovered by centrifugation as described in and single clone supernatants were prepared by infecting fresh *E.coli* K91K cells as described in section 3.3.

The single clone supernatants were used as templates in a PCR reaction identical to the one described in section 3.7. The PCR products were analysed by gel elctrophoresis (2% agarose) to confirm insertion of peptide encoding fragments. Furthermore the PCR products were purified (*QIAquick® PCR Purification Kit*, Qiagen) and shipped to GATC Biotech AB for sequencing using primer sekP (see 3.7). Sequencing of mutants AS-V3A and AS-F7A was carried out at the Laboratory for Immunohistochemistry and Immunopathology, Institute of Pathology, University of Oslo, Rikshospitalet.

3.10.3 Effect of amino acid substitutions on SC binding capacity

Single clone phage supernatants, confirmed to contain phage displayed alanine or serine substituted peptides, were screened for their ability to bind to immobilized SC *in vitro* by ELISA and elution-titration studies as described in section 3.5 and 3.6, respectively.

3.11 Construction of GFP fusion proteins

Auto fluorescent GFP fusion proteins were constructed to asses the pIgR binding properties of the phage displayed peptide C9-4 (CVVWMGFQQVC) when taken out of the gpIII fusion context. To ensure flexibility, the fusion proteins were constructed with a linker sequence. To examine the role of avidity on pIgR binding, a peptide-dimer fusion to GFP was designed. Another peptide, also selected by phage display, with affinity for complement subcomponent C1q was applied as control (CYWVGTWGEAVC). This peptide has previously been shown to retain target binding capacity when shifted from the phage context to alternative fusion partners (Lunde et al., 2002).

3.11.1 Production of GFP-fusion encoding inserts

The pHOG-GFP plasmid was propagated by transforming *E.coli* XL-1 blue cells as described in section 3.2. Single clones of the transformants were transferred to 3mL LB/amp growth medium in 15mL tubes and incubated at 37^{o}C ON with agitation. The following day pHOG-GFP was isolated from the ON cultures by *Wizard®Plus SV Miniprep* (Promega) as described in the protocol provided in the DNA isolation kit.

PCR was used to recover the GFP gene from the pHOG-GFP vector. Biotinylated primers, GFP-Start and GFP-End (Eurogenetic) were designed as described in table 2. GFP-Start carried an *Nco*I restriction site, whereas GFP-End carried an *Xho*I site, producing inserts compatible with directional cloning in pET21d (Novagen) (Figure 1A). Three fusion protein encoding fragments, carrying *Nco*I restriction sites, were constructed using specific, mutagenic oligonucleotides (DNA Technologies) (Table 2).

Table 2. GFP fusion primer sequences

Primer name	Oligonucleotide sequence *
GFP-Start	Bio-ATATATTATTCCATGGTGAGCAAGGGCGAGGAGCTG
GFP-End	Bio-ATAATAATACTCGAGTTACTTGTACAGCTCGTCCATGCCG
C9-4-Mono-GFP	ATAATATCCATGGGAGCCTGTGTGGTTTGGATGGGTTTTCAGCAGGTGTGT GGAGGACCAGGAGTGAGCAAGGGCGAGGAGCTG
C9-4-Dimer-GFP	ATAATATCCATGGGAGCCTGTGTGGTTTGGATGGGTTTTCAGCAGGTGTGT GGAGGAGGACCAGGATGTGTGGTTTGGATGGGTTTTCAGCAGGTGTGT GGAGGACCAGGAGTGAGCAAGGGCGAGGAGCTG
C1qC10-Mono-GFP	ATAATATCCATGGGAGCCTGTTATTGGGTGGGTACGTGGGGTGAGGCTGTTTGT GGAGGACCAGGAGTGAGCAAGGGCGAGGAGCTG

*All sequences are given in 5' - 3' direction.
Legend: - *Nco*I (CCATG) and *Xho*I (CTCGAG) restriction sites are highlighted in yellow
- GFP annealing primer regions are highlighted in turquoise.
- Sequences highlighted in red represent fusion peptide encoding DNA.
- A spacer region, connecting the two peptides of the dimer fusion peptide, is given in pink.
- The linker region, connecting the peptide to GFP, is given in blue.
- DNA highlighted in grey separates the restriction sites from the peptide encoding DNA. .

GFP was recovered by PCR as in the Vent polymerase based reaction described in 3.7 using primers GFP-Start (40pmol) and GFP-End (40pmol) and 10µL pHOG-GFP miniprep (10ng/µL) as template. Mutagenic primers (~40pmol) and GFP-End (40pmol) were employed to produce fusion protein encoding PCR products using 0.5µL *Herculase*^{TM} Enriched DNA polymerase (5000U/µL) and 10µL *Herculase*^{TM} reaction buffer (10x), in a reaction otherwise identical to the one described in 3.7.

The PCR products were analysed by gel electrophoresis (1% agarose), purified with *Centriflex*^{TM} protein removal cartridges, *seeDNA*^{TM} precipitated, resuspended in 20µL dH_2O and concentration determined by gel electrophoresis (1% agarose). 2µL of the purified DNA was restriction digested by 1µL *Xho*I (20000U/mL, NEB), 1µL *Nco*I (10000U/mL, NEB), 3µL 10x NEB2 buffer and dH_2O to a total volume of 30µL. Following incubation for 2h at 37^{o}C, the restriction reaction was analysed by gel electrophoresis (1% agarose).

Biotinylated ends and uncut DNA were removed from the fusion free GFP PCR product with *Streptavidine* as described in section 3.10.1. Restriction digested fusion protein inserts were *seeDNA*^{TM} precipitated and extracted from gels (1% agarose) using *QIAquick*® *Gel Extraction Kit* (Promega). Finally, the DNA inserts were *seeDNA*^{TM} precipitated and analysed by electrophoresis (2% *NuSieve*®3:1 agarose) to determine the DNA concentration.

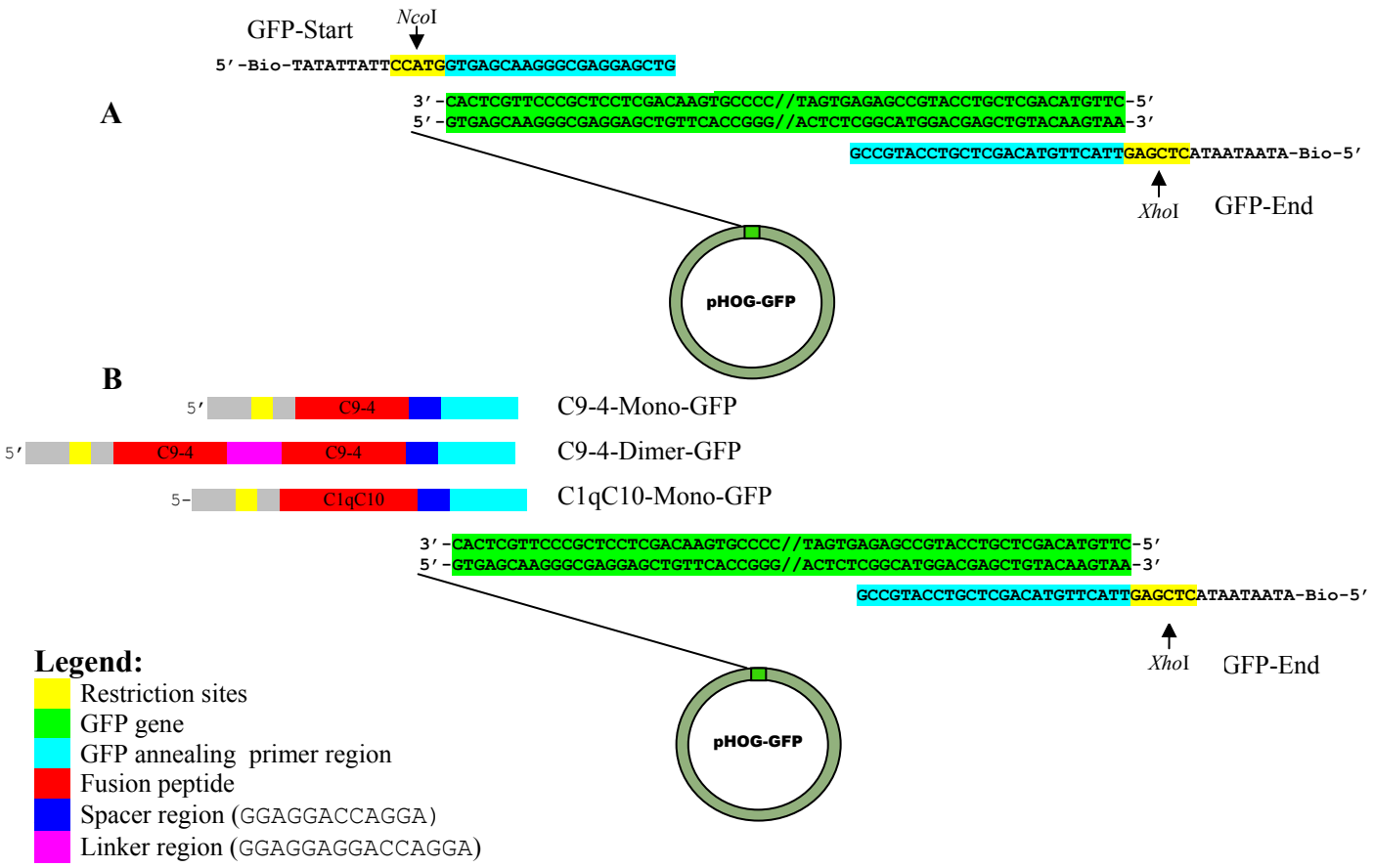


Figure 1. Overview of the GFP fusion protein cloning strategy
A-The GFP gene was recovered from a pHOG-GFP plasmid in a PCR, using biotinylated primers GFP-Start and GFP-End comprising *Nco*I and *Xho*I restriction sites respectively, to produce a pET21d compatible cloning insert.
B-GFP fusion protein encoding inserts were constructed using three different mutagenic primers carrying peptide encoding sequences (sequences given in table 2) and an *Nco*I restriction site. The PCR reactions were carried out using pHOG-GFP as template and the additional primer GFP-End to produce pET21d compatible fusion protein encoding inserts.

3.11.2 Ligation into pET21d, and transformation of *E.coli* BL21DE

10µL pET21d plasmid DNA was restricted, analysed, *seeDNA*^{TM} precipitated and resuspended in 20µL dH_2O as described in 3.11.1. Furthermore linearized vector DNA was purified using *QIAquick*® *Gel Extraction Kit*, *seeDNA*^{TM} precipitated and concentration determined by gel electrophoresis (2%*NuSieve*®3:1 agarose).

GFP encoding inserts with and without peptide fusions were ligated into linearized pET21d, using approximately 100ng of pET21d vector DNA and 1ng of the various inserts, as determined by gel electrophoresis band intensity, and *seeDNA*^{TM} precipitated as described in 3.10.2.

Electrocompetent *E.coli* BL21DE cells were transformed with pET21d carrying GFP and GFP fusion protein encoding inserts by electroporation as described in 3.2. Transformed single clones were transferred to 3mL LB/amp and incubated ON at 37°C with agitation. 850µL of the ON cultures were transferred to 1.5mL tubes containing 150µL Glycerol and stored at -70°C.

3.11.3 Plasmid isolation, restriction analysis and sequencing

pET21d DNA was isolated from the ON cultures using a *Wizard® Plus SV Miniprep* kit. 10µL plasmid aliquots were restriction digested with *Nco*I and *Xho*I as described in 3.11.1. The restriction reactions were analysed by gel electrophoresis (1% agarose) to confirm vector linearization and insert excision. 30µL plasmid prep was shipped to GATC Biotech AG for sequencing. The GFP-peptide encoding inserts were sequenced using primers T7 and pET-RP (GATC in-house primers).

3.11.4 Correction of deletion in peptide dimer fusion to GFP

Two guanine (G) deletions in the C9-4-Dimer-GFP construct were attempted corrected by design of an additional primer. The mega primer, C9-4-Dimer-G-correction-GFP (AGATATACCATGGGAGCCTGTGTGGTTTGGATGGGTTTTCAGCAGGTGTGTGGAGGAGGACCAGGATGTGTGGTTTGGATGGG) (DNA Technologies), was designed to anneal to the pET21d template carrying the deletions. The PCR reaction was carried out essentially as described in 3.11.1, with the following alterations: 10µL pET21d miniprep (~10ng/µL) carrying the deletions were used as template, 1µL of *Herculase™* DNA polymerase was added and the annealing temperature was adjusted to 60°C. The PCR product was prepared for cloning as described in 3.11.1 and 3.11.2.

3.12 Bacterial expression of GFP fusion proteins

Crude cell lysates were prepared from *E.coli* BL21DE cells, confirmed to be transformed with pET21d carrying the GFP constructs. GFP expression was verified by fluorescence microscopy, SDS PAGE and Western blotting. The feasibility of using ammonium sulphate precipitation as a purification method for bacterially produced GFP was investigated.

3.12.1 Bacterial cell growth and induction of protein expression

pET21d transformed *E.coli* BL21DE cells were plated onto ampicilline (50µg/mL) / glucose (100µL, 2M) plates and incubated at 37°C ON to yield fresh single colonies. Single colonies were transferred to 2mL ampicilline (100µg/mL) added LB growth medium and incubated for approximately 3h at 37°C with agitation. Subsequently the starter cultures were transferred to Erlenmeyer shaking flasks containing 100mL of ampicilline added (100µg/mL) LB and incubated at 37°C with agitation for good aeration until a density at A_{600} of ~1 was reached (PU 8750 UV/VIS scanning spectrophotometer). Pre-induction-samples of the cultures were recovered (T_0). Protein expression was induced by adding 100µL IPTG (100mM) to the cultures and the incubation at 37°C was continued for three hours (T_{180}). The cells were harvested by centrifugation at 4600rpm for 10 minutes at 4°C (Centra CL3, International Equipment Company) and resuspended in 10mL PBS per 100mL culture volume.

3.12.2 Fluorescence microscopy and cell lysis

Auto fluorescent GFP was detected by fluorescence microscopy (Axioplan 2 Imaging, Zeiss). Resuspended cells were frozen and thawn prior to ultrasonic disruption in 50mL tubes on ice for 4x 10 seconds at max power (Soniprep 150, MSE Scientific Instruments). Cell debris and unlyzed cells were removed by centrifugation 2x 30min at 4600rpm (Centra CL3, International Equipment Company). The lysates were stored at -70°C.

3.12.3 SDS PAGE

Cells and crude lysates were analyzed by SDS PAGE. 12µL of the samples to be analysed and dilutions of GFP (BD Biosciences) were mixed with 4µL of 4xSDS loading buffer, incubated for 5 minutes at 95°C and loaded on to 12% Tris-HCl pre-casted gels (BioRad). Molecular weight standards *Rainbow*™ (Amersham Biosciences) and SDS PAGE broad range standard (BioRad) were used.

Following electrophoresis (BioRad *Mini PROTEAN® II Cell*, BioRad powerpac, 90-100V), the gels were *Coomassie*™ (PhastGel Blue R., Amersham Pharmacia Biotech) stained for 15minutes, fixated and washed ON in a methanol (30%), acetic acid (10%) added fixing solution. Images of the stained gels were obtained using a BioDoc-It™ UV Transluminator (UVP).

3.12.4 Western Blotting

GFP in the lysates diluted 1:10 in PBS and dilutions of commercially available GFP were detected in a western blott alongside biotinylated SDS PAGE broad range molecular weight standard (BioRad). The proteins were separated as described 3.12.3 and electro transferred to an *Immobilon*^{TM}-P PVDF membrane (Millipore) by blotting at 4^{o}C for 1hour (BioRad minigel blotting cassette, BioRad powerpac, 100V). The membrane was allowed to air dry for a minute prior to blocking in 1% SKM, at 4^{o}C, ON in a rotating 50mL tube. The membrane blocking was continued for 30min at RT. Subsequently, the membrane was washed thoroughly in PBS/T. The membrane was incubated with goat polyclonal biotinylated antibody against GFP (1mg/mL) (Abcam) diluted 1:2000 in PBS/T for 1hour at RT. The washing step was repeated before the membrane was incubated with *Streptavidine*-HRP conjugate (Amersham LIFE SCIENCES) diluted 1:5000 in PBS/T for 1hour at RT. Following repeated washing, the membrane was incubated in *ECL*^{TM} western blotting detection reagent (Amersham Pharmacia Biotech) for 1minute at RT. Biomax MR- film (Kodak) was exposed to the membrane and developed using an Optimax processor (Protec).

3.12.5 Ammonium sulphate precipitation

Cell lysate from *E.coli* BL21DE cells expressing GFP without fusion peptide (GFP) was ammonium sulphate precipitated. Increasing amounts of ammonium sulphate ($(NH_4)_2SO_4$, Merck), 5, 10, 20, 30, 40 and 50 % (1g/100mL = 1%), were added to 1.5 mL aliquots of lysate. The samples were kept cool and agitated until all of the salt had dissolved. The samples were incubated for one hour on ice prior to centrifugation at 10000rpm for 30min at 4^{o}C (Biofuge A, Heraeus sepatech). Supernatants were discarded and the pellets were dissolved in 1mL PBS. Subsequently the samples were analysed by SDS PAGE as described in 3.12.3.

3.13 Target binding capacity of GFP fusion proteins

Crude cell lysates, containing GFP and GFP fusion proteins, were screened for their ability to bind immobilized SC and C1q by direct measurement of fluorescence and by ELISA. The lysates were standardized by measurement of the relative fluorescence (485nm/535nm, 1sec, *Victor*^{TM} fluorometer, Perkin Elmer) in black *FluoroNunc*^{TM} *Maxisorp*^{TM} plates and subsequent dilution in PBS to ensure consistent in-put levels of GFP.

3.13.1 Target binding determined by fluorescence measurement

Black *FluoroNunc*^{TM} *Maxisorp*^{TM} plates were coated with 100µL SC (10µg/mL) or 100µL C1q (10µg/mL) (Calbiochem) in separate wells by ON incubation at 4^{o}C. Following 1hour of blocking in 1% SKM, dilutions of lysates in PBS (GFP-no fusion, C9-4-Mono-GFP, C1qC10-Mono-GFP and GFP negative lysate) were added to the wells and incubated for 90minutes. Subsequently, the wells were washed in PBS/T (5x using a Tecan Columbus washer, Tecan). 100µL of PBS was added to the wells prior to fluorescence measurement to confirm target binding (485nm/535nm, 1sec, *Victor*^{TM} fluorometer, Perkin Elmer). Dilutions of commercially available GFP (BD Biosciences) were measured alongside the lysates for comparison and concentration determination.

3.13.2 Target binding determined by ELISA

Nunc^{TM}*Maxisorp*^{TM} plates were coated with SC and C1q as described in 3.13.1, blocked in 1% SKM and washed as described 3.5. Standardized lysates (~20000RFU, corresponding to ~10µg/mL of commercially obtained GFP) were pre-incubated (1:1) in blocking solution prior to 90minutes incubation with the target at RT. Following washing, as described in 3.5, 100µL of biotinylated goat polyclonal anti-GFP antibody (Abcam) diluted 1:1000 in blocking solution was added to the wells and incubated for 1h at RT. The washing procedure was repeated. Next, 100µL of *Streptavidine*-HRP conjugate (Amersham LIFE SCIENCES) diluted 1:4000 in blocking solution was added and incubated for 1h at RT. Following repeated washing, bound GFP fusions were visualized by adding ABTS. Absorbance at 405nm (A_{405}) were measured (Tecan Sunrise Spectrophotometer, with Magellan 3.0 software) after 30-60 minutes incubation at RT. The integrity of the SC coat was confirmed by the use of a SC binding phage detected essentially as described in 3.5. The C1q coat quality was assessed using a rabbit anti-C1q antibody (3.7mg/mL, DAKO) diluted 1:1000 in 1% SKM and an anti-rabbit antibody conjugated to HRP (Amersham Biosciences) diluted 1:4000 in 1% SKM essentially as in the lysate ELISA described above.

4. Results

The objective of this work has been to screen for, and isolate phage displayed peptides with affinity for SC and potential for use in mucosal targeting. A semi-automated screening procedure has been developed to increase sample throughput and allow for efficient screening of large numbers of phage clones.

4.1 Affinity selection

To isolate novel peptides with high affinity for human SC, two rounds of affinity selection from the random phage display libraries Cys9 and 7-mer, were performed. To reduce unspecific background binding, the solid phase was blocked, and phages were pre-incubated, using 1% SKM or 1% BSA.

1.5 x 10^9 TUs of the 7-mer library were added to wells of *Nunc-Immuno*^{TM} *MaxiSorp*^{TM} strips coated with 200µL of SC at 30µg/mL. The 7-mer library contained approximately 8 x 10^7 different clones. Thus, on average only ~20 particles of each clone were represented in the first round of affinity selection. To the second round of selection 1.0 x 10^9 TUs from the 7-mer amplified eluate 1 obtained using BSA as blocking agent (7-mer EA1BSA) were added. The same amount of phage from the first 7-mer amplified eluate, obtained using SKM (7-mer EA1SKM), were added to the second selection round.

1.4 x 10^{10} TUs of the Cys9 library were added to *Nunc-Immuno*^{TM} *MaxiSorp*^{TM} tubes coated with 800µL of SC at 30µg/mL. The library contained roughly 5 x 10^7 different clones, thus ~280 copies of each clone was represented in the first round of selection. 4.5 x 10^9 TUs from the Cys9 amplified eluate 1, obtained using BSA as blocking agent, (Cys9EA1BSA) was added to the second round of selection. 3.0 x 10^9 TUs from the Cys9 amplified eluate 1, obtained using SKM as blocking agent (Cys9EA1SKM), were added to the second round of selection using SKM as blocking agent.

Recovery of 7-mer phage from the first- and second round of selection using BSA as blocking agent (7-mer E1BSA) 7.7 x 10^{-3} % and (7-mer E2BSA) 6.0 x 10^{-3} %, respectively, gave no indication of enrichment for SC binding phages. Recovery of 7-mer phage from the first- and second round of selection using SKM as blocking agent (7-mer E1SKM) 1,2 x 10^{-3} % and (7-mer E2SKM) 2.1 x 10^{-4} % , respectively showed a slight increase of phage yield.

After the first- and second round of selection using the Cys9 library with BSA as blocking agent (Cys9 E1BSA) and (Cys9 E2BSA) recovery rates of 2.1 x 10^{-3} % and 1.2 x 10^{-3} %, respectively, were seen. SKM blocking after the first- (Cys9 E1SKM) and second (Cys9 E2SKM) round returned 2.3 x 10^{-3} % and 6.0 x 10^{-4} % phage yield, respectively. Thus, no enrichment of SC binding phages was seen after two rounds of selection using Cys9 regardless of the blocking agent applied. The results after two rounds of affinity selection are summarized in table 3.

Table 3. Affinity selection for SC binding peptides phage display libraries

Library. eluate and blocking agent	Input (TU) [a]	Output (TU)[b]	Yield (%)[C]
7-mer E1 BSA	1.5 x 10^{9}	115000	7.7 x 10^{-3}
7-mer E2 BSA	1.0 x 10^{9}	60000	6.0 x 10^{-3}
7-mer E1 SKM	1.5 x 10^{9}	3200	2.1 x 10^{-4}
7-mer E2 SKM	1.0 x 10^{9}	12000	1.2 x 10^{-3}
Cys9 E1 BSA	1.4 x 10^{10}	300000	2.1 x 10^{-3}
Cys9 E2 BSA	4.5 x 10^{9}	52000	1.2 x 10^{-3}
Cys9 E1 SKM	1.4 x 10^{10}	320000	2.3 x 10^{-3}
Cys9 E2 SKM	3.0 x 10^{9}	18000	6.0 x 10^{-4}

a) 10µL of the 7-mer library (titer 1.5x10^{11}TU/mL) and 100µL of the Cys9 library (titer 1.4x10^{11}TU/mL) served as input to the first round of affinity selection. 100µL of 7-mer EA1BSA/SKM (titre 1.0x10^{10}TU/mL) served as input to the second panning round as did 100µL of Cys9EA1BSA (titre 3.0x10^{10}TU/mL) and Cys9EA1SKM (titre 4.5x10^{10}TU/mL).

b) Output phages were calculated as the total number of *E.coli* K91K TUs in the eluates.

c) Yield = (output TU/input TU) x 100%

4.1.1 Screening the amplified eluates for SC binding phages by ELISA

To disclose whether the pool of amplified phages after the second round of selection contained detectable amounts of SC binding phages, an ELISA was performed. *Nunc-Immuno*^{TM} *MaxiSorp*^{TM} 96 well plates coated with 100µL of SC at 5µg/mL were used as solid phase and 50µL of each amplified eluate was added to the wells diluted 1:1 in blocking solution. The phage clones C9-4 and T7 were used as positive and irrelevant control phages, respectively.

The amplified eluates from the 7-mer library selection, namely 7-mer EA1BSA, 7-mer EA2BSA, 7-mer EA1SKM and 7-mer EA2SKM, were assayed for SC binding and binding to blocking solutions or the solid phase. No difference was seen between binding to wells with- or without SC coat. Thus, despite the increased recovery of phage that was observed between rounds of selection, using SKM as blocking agent, none of the phage pools contained detectable amounts of SC binding phage (data not shown).

Likewise, the phage pools obtained after Cys9 library selection, namely Cys9EA1BSA, Cys9EA2BSA, Cys9EA1SKM and Cys9EA2SKM, were assayed for SC binding and binding to blocking solutions or the solid phase. The ELISA assay included cross-blocking where eluates obtained using a SKM blocking procedure were blocked using BSA and vice versa to disclose the role of the blocking agent. From wells coated with SC and blocked with SKM a small increase in readout value A_{405} between Cys9EA1SKM and Cys9EA2SKM was obtained. No such increase in binding to SKM blocked wells alone was seen, indicating the presence of SC binding phages in the pool. The Cys9EA2SKM A_{405} value was clearly above background level. When the same amplified eluates were screened for SC binding in wells blocked with BSA, the same trend occurred only this time the absorbance values were higher and both Cys9EA1SKM and Cys9EA2SKM values exceeded the background level.

An increase in absorbance values was also seen between Cys9EA1BSA and Cys9EA2BSA. However, the same increase was seen for phage binding to BSA blocked wells as well. When subjected to cross blocking with SKM, the increase in A_{405} value between Cys9EA1BSA and Cys9EA2BSA is preserved. The results of the Cys9 phage pool ELISA are summarized in figure 2.

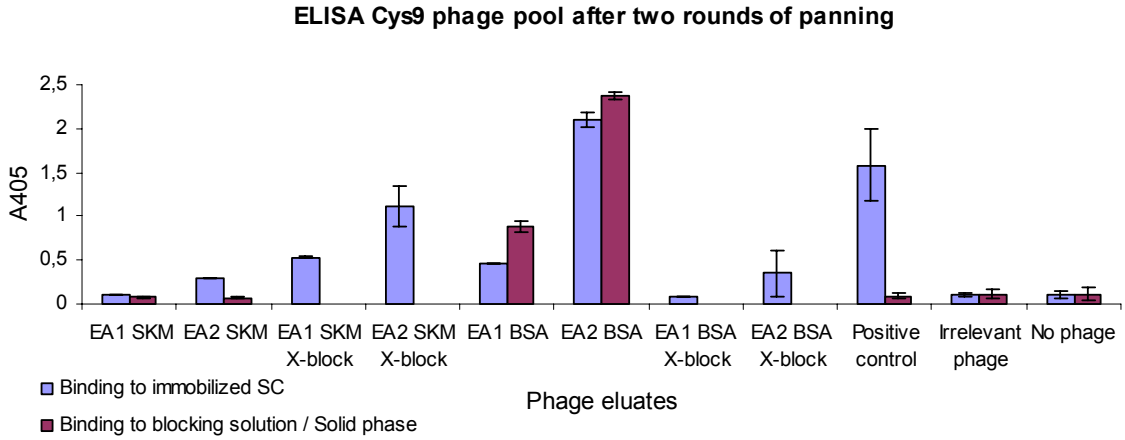


Figure 2. Cys9 phage pool SC binding and enrichment after two rounds of panning
Amplified eluates after two rounds of panning were screened for SC binding (light blue) and background binding (purple) detected in an ELISA as absorbance at 405nm. Furthermore, the role of the blocking agent in the ELISA was revealed in an x-blocking assay. Background binding was not determined in the x-blocking experiment. *Nunc-Immuno*^{TM} *MaxiSorp*^{TM} 96 well plates coated with 100µL of SC at 5µg/mL were used as solid phase.

The ELISA results indicate that Cys9EA1SKM and Cys9EA2SKM contain detectable amounts of SC binding phages.

4.2 Screening for single clones displaying SC binding peptides

To screen for single clones expressing SC binding peptides, randomly picked clones originating from *E.coli* K91K cells transduced with phage from all four Cys9 amplified eluates were grown in 96 deep well plates compatible with semi-automated sample handling. Single clone phage supernatants were screened for SC binding and background binding by ELISA to identify those specifically binding to SC. SC specific clones, selected by A_{405} readout values, were further investigated in subsequent ELISAs. Finally, the sequences of the SC binding peptides were determined.

4.2.1 Initial high throughput screening for SC binding single clone phages

Scripts for use of the automated liquid handler, Biomek2000, were designed to allow for efficient ELISA screening (see Appendix 1). The semi-automated procedure was never implemented due to technical difficulties with the robotic washing tool. 88 transductants originating from Cys9EA1SKM Cys9EA2SKM, Cys9EA1BSA and Cys9EA2BSA (88 x 4=352 clones) were propagated in 96 deep well plates. *Nunc-Immuno*^{TM} *MaxiSorp*^{TM} 96 well plates coated with 100µL of SC at 5µg/mL were used as solid phase and 100µL of each single clone supernatant diluted 1:1 in blocking solution was added to the wells. The phage clones C9-4 and T7 were used as positive and irrelevant control phages, respectively.

6 of the 88 clones from Cys9EA1SKM were positive for SC binding. As many as 75 clones out of the 88 from Cys9EA2SKM were also able to bind immobilized SC specifically. 63 of the 88 Cys9EA1BSA clones appeared to be positive for SC binding. However, all came back as non-specific binders. 9 out of the 88 Cys9EA2BSA clones showed specific SC binding, whereas 68 clones proved to be background binders in the primary binding assay (data not shown). The results of the initial screening for SC binding single clone phages are summarized in table 4.

Table 4. Initial screening for SC binding single clone phages

EA of origin	Number of screened clones	Number of SC binding clones	Number of background binding clones
EA1SKM	88	6	0
EA2SKM	88	75	0
EA1BSA	88	0	63
EA2BSA	88	9	68

4.2.2 Screening of selected positive individual clones in a confirmative ELISA

To confirm the SC binding capacity of selected single clones, identified in the large scale screening described above, additional ELISA assays were performed. The confirmative ELISAs were identical to the initial high throughput assay.

All 6 of the identified EA1SKM single clones were confirmed to bind SC. No background binding was identified (Figure 3A). Furthermore, of the 75 EA2SKM positives initially identified, the ten clones with the highest A_{405} values in the ELISA were picked, accompanied by 20 more with varying A_{405} readout values. All of the top ten A_{405} clones and 18 of the other 20 clones came back positive for SC binding in the confirmation ELISA (Figure 3B). Also, all of the 9 selected EA2BSA single clones returned A_{405} values above the negative control. One of the clones, 6C, gave high readout values for background binding and was not sequenced (Figure 3C).

The results of the confirmative single clone ELISAs are recapitulated in figure 3.

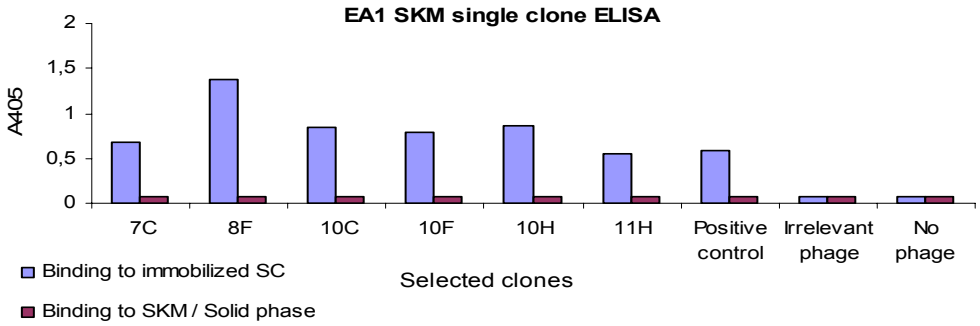


Figure 3A. EA1SKM single clone binding to immobilized SC as determined by ELISA
Single clones, obtained from host cells infected with EA1SKM, were examined for their ability to bind immobilized SC. 6 out of 88 clones proved to bind SC specifically. The clones are named after their matrix position in the 96-well growth plate in which they were propagated. Binding to SC and background binding is depicted in light blue and in purple, respectively.
Note: This ELISA was prepared with single samples of each clone only, and it has not been repeated.

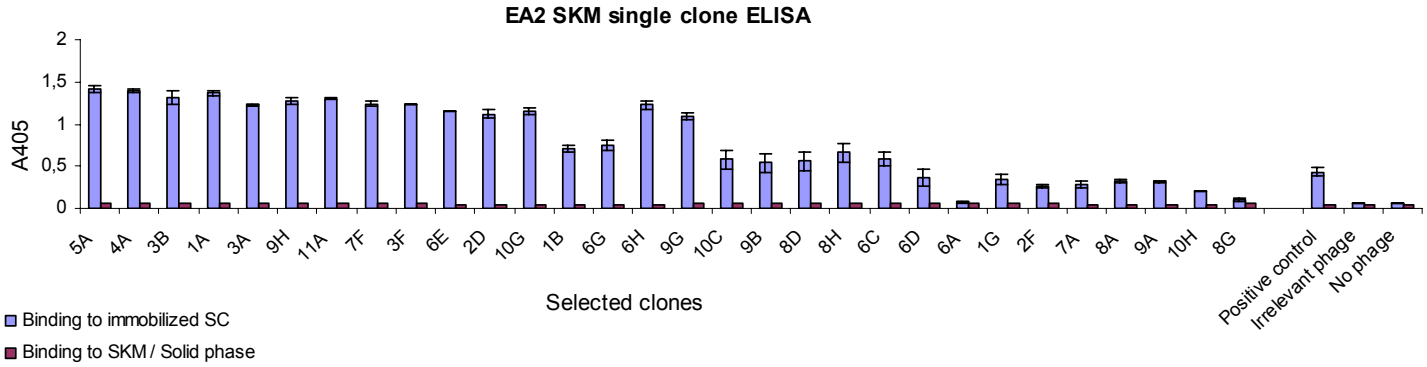


Figure 3B. EA2SKM single clone binding to immobilized SC as determined by ELISA
Single clones, obtained from host cells infected with EA2SKM, were examined for their ability to bind immobilized SC. 10 clones selected for their top ten A_{405} readout values in an initial ELISA (5A-6E) and 20 clones (2D-8G) with varying A_{405} readout value in the same experiment were selected to undergo further investigation. The clones are named after their matrix position in the 96-well growth plate in which they were propagated. Binding to immobilized SC and background binding, depicted in light blue and in purple, respectively, were determined with samples in triplicates.

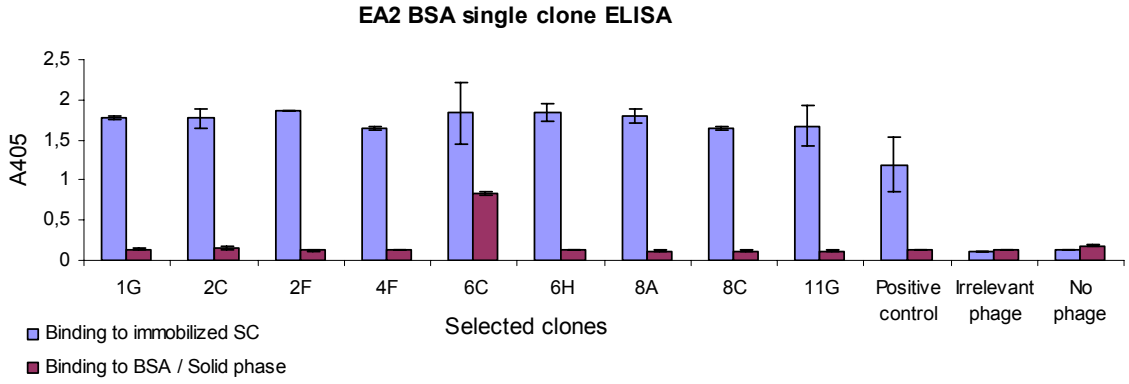


Figure 3C. EA2BSA single clone binding to immobilized SC as determined by ELISA
Single clones, obtained from host cells infected with EA2BSA, were examined for their ability to bind immobilized SC. 9 clones selected for their apparent SC binding in an initial ELISA were selected to undergo further investigation. The clones are named after their matrix position in the 96-well growth plate in which they were propagated. Binding to immobilized SC and background binding, depicted in light blue and in purple respectively, were determined with samples in duplicates.

4.2.3 Sequencing of selected SC binding phage

To produce sequencing templates of the DNA encoding the peptides fused to gpIII, supernatants from selected, ELISA positive, single clones were used as template in a PCR. A total of 44 sequencing templates from Cys9EA1SKM (6), Cys9EA2SKM (30) and Cys9EA2BSA (8) were shipped to GATC Biotech AG, and 43 sequences were obtained (all but Cys9EA2SKM clone 10G).

7 different sequences were isolated in the screening procedure. Only two of these encoded cysteine constrained 9-mer peptides. 6 clones displayed peptide CIIVPHAYAWC, all of which originated from Cys9EA2SKM. 13 clones with sequence CVVWMGFQQVC were identified, one from Cys9EA1 SKM and 12 from Cys9EA2 SKM. This peptide is identical to the one displayed by the positive control phage C9-4 that had previously been selected and characterized as a SC binder (Braathen et al., unpublished).

The sequence CALVSEAGCLVWAA was found to be displayed by a total of 20 selected single clones, including all 8 Cys9EA2 BSA clones, 4 Cys9EA1 SKM clones and 8 Cys9EA2 SKM clones. This dominating 14-mer peptide sequence is in-frame without cysteine (TGT) as the C-terminal amino acid. Thus, it deviates from the expected cysteine flanked random 9-mer. In the following, this peptide/phage is referred to as ASR.

Four 7-mer peptides were also identified. Three of them were isolated from phage originating from Cys9EA2SKM (RGPFVYM, RFWWGWY and GWAGWLG), whereas the last one (PFVVLLV) was obtained from an EA1 SKM clone. It should be noted that Cys9EA2SKM clones 6A (RGPFVYM) and 8G (GWAGWLG) tested negative for SC binding in the confirmation ELISA. However, at that time they were already shipped of to sequencing. The recovered 7-mers are likely to be contaminants from the affinity selection performed in parallel to the Cys9 library using a 7-mer library. ASR and C9-4, represented approximately 45% and 30% of the recovered sequences, respectively.

Abi files, representing all of the 7 recovered peptides, describing the detailed sequencing data are enclosed in appendix 2. The single clone phage sequencing results are summarized in table 5.

Table 5. Recovered sequences from SC affinity selection

DNA sequence	Amino acid sequence	Eluate of origin, selected clone and frequency *
TGT GCG TTG GTG AGT GAG GCT GGC TGC TTA G**TG T**GG GCC GCT	CALVSEAGCLVWAA	EA2 BSA **1G, 2C, 2F, 4F, 6H, 8A, 8C, 11G** 8/8 EA1 SKM **7C,10C, 10F, 10H** 4/6 EA2 SKM **1B, 6G, 10C, 9B, 8D, 8H, 6C, 3F** 8/29
TGT GTG GTT TGG ATG GGT TTT CAG CAG GTG TGT	CVVWMGFQQVC	EA1 SKM **8F** 1/6 EA2 SKM **5A, 4A, 3B, 1A, 3A, 9H, 11A, 7F, 6E, 2D, 6H, 9G** 12/29
TGT ATT ATT GTG CCG CAT GCG TAT GCT TGG TGT	CIIVPHAYAWC	EA2 SKM **6D, 1G, 2F, 7A, 8A, 9A** 6/29
CGT GGT CCG TTT GTG TAT ATG	RGPFVYM	EA2 SKM **6A** 1/29
AGG TTT TGG TGG GGG TGG TAT	RFWWGWY	EA2 SKM **10H** 1/29
GGG TGG GCT GTG TGG CTG GGG	GWAGWLG	EA2 SKM **8G** 1/29
CCG TTT GTG GTG TTG TGG GTG	PFVVLLV	EA1 SKM **11H** 1/6

* Cys9 eluates from which the sequences are derived, each individual clone representing this sequence (**bold**) and finally, the frequency of each recovered sequence in relation to their eluate of origin.

4.3 Inhibition of phage binding to SC by pIgA, pIgM and SpsA fragments

C9-4 and ASR were tested for interaction with immobilized SC in a competition ELISA in the presence of IgA, IgM or SpsA. *Nunc-Immuno*^{TM} *MaxiSorp*^{TM} 96 well plates coated with 50µL of SC at 30µg/mL were used as solid phase. Phage supernatant was added to the wells diluted 1:1 in dilutions of the inhibitors.

C9-4 binding to immobilized SC is not inhibited by pIgA or pIgM. SpsA fragment SH2 and SM1 on the other hand, knock out detectable C9-4 binding to immobilized SC (see figure 4A). A SC-negative control fragment of SpsA, SH3, did not interfere with the C9-4 SC interaction. In conclusion, C9-4 binds SC at a location different from the binding site used by pIgA and pIgM. Moreover, the interaction site is likely to overlap a contact area also used by SpsA. ASR binding to SC is not affected by the polymeric Igs or by SC binding SpsA fragments. Thus, ASR is likely to bind an area of the receptor not utilized by any of the other ligands applied (see figure 4B). The results of the inhibition study are summed up in figure 4.

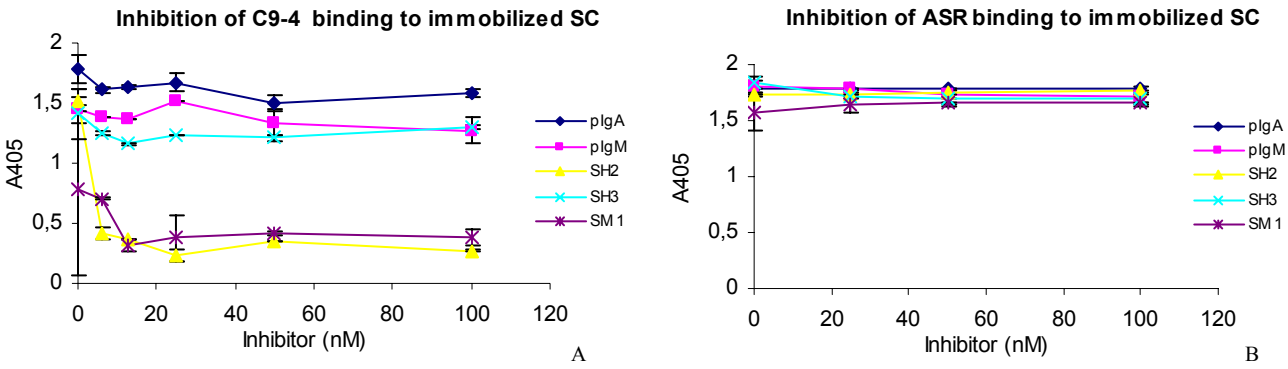


Figure 4. Inhibition of phage binding to immobilized SC
A: pIgA (blue) does not inhibit C9-4 binding to hSC, neither does pIgM (pink) nor SpsA fragment SH3 (turquoise). SH2 (yellow) and SM1 (purple), on the other hand, clearly inhibit C9-4 binding to the SC coat. B: None of the added inhibitors hamper the interaction between ASR and immobilized SC.

4.4 Transcytosis of SC binding phage across pIgR transfected MDCK cells

To elucidate whether C9-4 and ASR retain binding to SC when the target is presented in its native configuration and to reveal if the phage can be transported across epithelia by pIgR mediated transcytosis, phages were allowed to react with polarized MDCK cells transfected with pIgR. Untransfected cells were applied as control. Motivated by the inhibition studies, the ability of both C9-4 and ASR phages to discriminate between murine and human SC was investigated by using MDCK cells transfected with murine or human pIgR in the transcytosis experiments. Basolateral to apical transport was measured as phage recovered apically. Phages with irrelevant specificity (TCys9) as well as IgG were used to control for unspecific transport and leakage.

Basolateral inputs of C9-4, ASR and control phage were 8.2×10^{9} TU, 4.4×10^{8} TU and 2.7×10^{10} TU, respectively. 5×10^{-2} % of C9-4 was recovered apically of MDCK cells transfected with human pIgR. Cells transfected with the murine receptor returned only 2×10^{-5} %, the same as untransfected cells. In conclusion, it appears that C9-4 is transported across MDCK cell monolayers by pIgR mediated transcytosis. Evidently C9-4 also discriminates between murine and human pIgR, showing affinity for the human receptor only as SpsA does. Thus, C9-4 meet valuable criteria for peptide candidates intended for use in pIgR mediated mucosal targeting.

MDCK cells transfected with the human receptor, returned 4.7 x 10^{-4} % of the added ASR phage. Cells transfected with the murine receptor returned 7.2 x 10^{-4} %, whereas untransfected cells returned 5.4 x 10^{-4} % ASR phages. In conclusion, the ASR phage is not transported across the epithelial model system by pIgR mediated transcytosis. Whether this is due to lack of receptor binding on the cell surface or not has not been established.

The recovery figures for the irrelevant phage were as follows: 2.9 x 10^{-4} % for the human receptor; 1.2 x 10^{-4} % for the murine receptor and 1.3 x 10^{-3} % for untransfected cells. 15µg/mL IgG at 100nM was added in the basolateral medium. Less than ~30ng/mL IgG was recovered apically, confirming no system leakage (data not shown).
The results of the transcytosis experiment are summarized in figure 5.

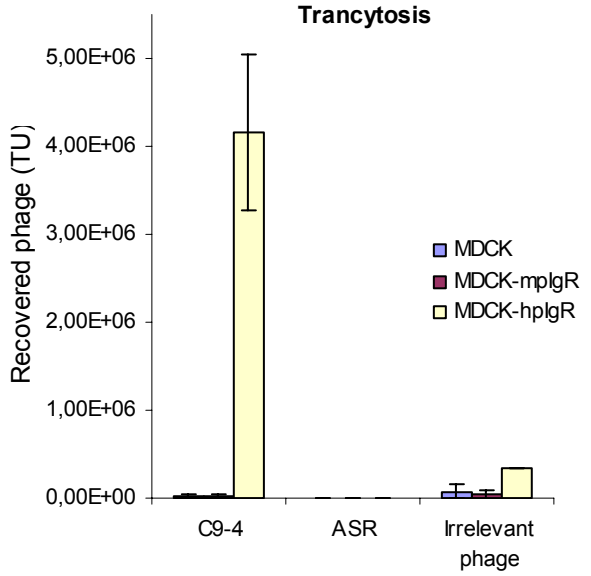


Figure 5. Transcytosis across MDCK cells

C9-4 phages (left) are specifically trancytosed across MDCK cells transfected with human pIgR (light yellow column). Untransfected cells (blue column) or cells bearing the murine receptor (purple column) are not capable of relocating the phage from the basolateral- to the apical side. ASR (centre) and the irrelevant control phage (right) are not able to utelize pIgR mediated transcytosis for transport. Basolateral inputs of C9-4, ASR and control phage were 8,2x10^{9} TU, 4,4x10^{8} TU and 2,7x10^{10} TU, respectively.

Apical medium was harvested and titrated as *E.coli* K91K TUs to calculate the yield of recovered phages.

4.5 Determination of the contribution of each C9-4 amino acid on SC binding

To determine the contribution of each individual amino acid on the functionality of the peptide each amino acid of C9-4 was systematically substituted with alanine. The flanking cysteines of the constrained peptide were substituted with serine. The mutants were sequenced to confirm substitutions before SC binding capacity was assessed by ELISA and elution-titration.

4.5.1 Sequencing of alanine/serine substitution mutated C9-4

Sequencing of the mutated DNA, confirmed the presence of the alanine/serine substitutions in the appropriate positions. The sequencing results are summarized in table 6. Abi files, representing C9-4 and all of the 12 substituted peptides, describing the detailed sequencing data are enclosed in appendix 2 and 3, respectively.

Table 6. Confirmed alanine / serine substitution sequences

Mutant	Peptide encoding inserts	Amino acid sequence
C9-4 *	TGT GTG GTT TGG ATG GGT TTT CAG CAG GTG TGT	C V V W M G F Q Q V C
AS-C1S	AGC GTG GTT TGG ATG GGT TTT CAG CAG GTG TGT	S V V W M G F Q Q V C
AS-V2A	TGT GCA GTT TGG ATG GGT TTT CAG CAG GTG TGT	C A V W M G F Q Q V C
AS-V3A	TGT GTG GCA TGG ATG GGT TTT CAG CAG GTG TGT	C V A W M G F Q Q V C
AS-W4A	TGT GTG GTT GCA ATG GGT TTT CAG CAG GTG TGT	C V V A M G F Q Q V C
AS-M5A	TGT GTG GTT TGG GCA GGT TTT CAG CAG GTG TGT	C V V W A G F Q Q V C
AS-G6A	TGT GTG GTT TGG ATG GCA TTT CAG CAG GTG TGT	C V V W M A F Q Q V C
AS-F7A	TGT GTG GTT TGG ATG GGT GCA CAG CAG GTG TGT	C V V W M G A Q Q V C
AS-Q8A	TGT GTG GTT TGG ATG GGT TTT GCA CAG GTG TGT	C V V W M G F A Q V C
AS-Q9A	TGT GTG GTT TGG ATG GGT TTT CAG GCA GTG TGT	C V V W M G F Q A V C
AS-V10A	TGT GTG GTT TGG ATG GGT TTT CAG CAG GCA TGT	C V V W M G F Q Q A C
AS-C11S	TGT GTG GTT TGG ATG GGT TTT CAG CAG GTG AGC	C V V W M G F Q Q V S
AS-C1SC11S	AGC GTG GTT TGG ATG GGT TTT CAG CAG GTG AGC	S V V W M G F Q Q V S

*C9-4 is included in the table for comparison only.

4.5.2 Effect of amino acid substitutions on binding to SC as determined by ELISA

To examine the effect of systematic alanine substitution and removal of flanking cysteines on target binding capacity, phages expressing mutated peptides were screened for SC binding using ELISA. C9-4 was used as positive control and TCys9, with irrelevant specificity, was used as negative control. *Nunc-Immuno*^{TM} *MaxiSorp*^{TM} 96 well plates coated with 200µL of SC at 5µg/mL, were used as solid phase and 100µL of each mutant supernatant was added to the wells diluted 1:1 in blocking solution. The amount of input phage in the ELISA varied from 2 x 10^9 TU to 7 x 10^9 TU. The lowest supernatant titre was seen for AS-C11S, whereas the highest was seen for the double mutant, AS-C1SC11S.

The ELISA reveals that all, but three, amino acids are essential to C9-4's ability to bind SC. Mutants AS-Q9A, and to a lesser extent AS-V3A and AS-G6A, bind SC. Substitutions of valine (V), glycine (G) and glutamine (Q) in positions 3, 6 and 9, respectively, are tolerated and SC binding are detected by ELISA. In conclusion, glutamine in position 9, valine in position 3 and glycine in position 6 appears not to be indispensable to the functionality of C9-4. The variation in phage input, does not explain the observed binding pattern. ELISA results from the substitution mutant binding to SC are summarized in figure 6.

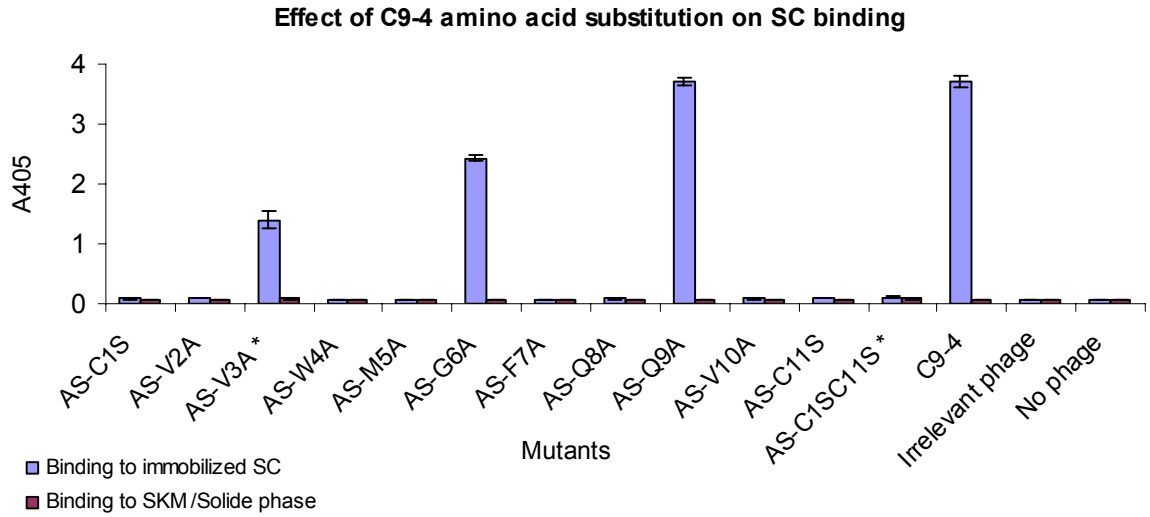


Figure 6. Amino acid substituted C9-4 peptide mutant binding to immobilized SC
The C9-4 substitution mutant phage presented peptides were screened in an ELISA to disclose amino acids essential to C9-4's SC binding capacity. Mutants AS-Q9A, and to a lesser extent AS-V3A and AS-G6A, bind SC. Valine (V), glycine (G) and glutamine (Q) in positions 3, 6 and 9, respectively, may be substituted. The remaining amino acids, cysteines included, appear to be critical to for SC binding. SC binding and binding to SKM or the solid phase are given in light blue and purple, respectively.

* AS-V3A and AS-C1SC11S were examined for SC binding in separate experiments with internal controls reaching the same A_{405} levels as shown here.

4.5.3 Effect of amino acid substitution on SC binding as determined by elution-titration

ELISA and elution-titration studies are alternative methods to investigate binding of phage clones to immobilized target. An elution-titration experiment was carried out to determine the number of phages added and recovered from each well. The yield of eluted phages was calculated, to determine the contribution of each individual amino acid.

Nunc-Immuno^{TM} *MaxiSorp*^{TM} 96 well plates coated with 200µL of SC at 20µg/mL were used as solid phase and 50µL of each mutant supernatant was diluted to a total volume of 200µL in SKM and added to the wells. Included controls were as in the ELISA described above. The amount of input phage to the elution-titration assay varied from 1 x 10^9 TU to 3.5 x 10^9 TU.

AS-Q9A clearly binds immobilized SC. The yield of AS-Q9A phage recovered from the procedure is approximately 60% of that seen for the positive control. As for the ELISA, AS-V3A and AS-G6A stand out from the remaining mutants.

Although AS-V3A and AS-G6A recovery only make up approximately 1,5% and 2% of the positive control, respectively, they return ~30-100 times more phage that do the other mutants and ~500-600 times more phage than do the irrelevant control phage. The recovery rate for all mutants, even the ones considered to be SC binding negatives (1.5×10^{-4} % - 9.0×10^{-5} % recovery), were higher than the number seen for the irrelevant phage (7.5×10^{-6} % recovery), this matter of fact was not revealed by ELISA. The results of the elution-titration study are summarized in table 7.

In conclusion, glutamine in position 9, valine in position 3 and glycine in position 6 appears not to be of critical importance to the functionality of C9-4. The elution and titration assay, confirms the data obtained in the ELISA and gives an estimate of the actual amount of phages bound to the target, thus the method return high resolution data.

Table 7. Effect of C9-4 amino acid substitution on SC binding as determined by elution / titration

Phage	Titre (TU/mL)	Input (TU)	Output (TU)	Yield (%)
AS-C1S	4.0×10^{10}	2.0×10^{9}	600	3.0×10^{-5}
AS-V2A	4.0×10^{10}	2.0×10^{9}	1300	6.5×10^{-5}
AS-V3A*	5.0×10^{10}	2.5×10^{9}	75000	3.0×10^{-3}
AS-W4A	5.0×10^{10}	2.5×10^{9}	900	3.6×10^{-5}
AS-M5A	5.0×10^{10}	2.5×10^{9}	3750	1.5×10^{-4}
AS-G6A	4.0×10^{10}	2.0×10^{9}	80000	4.0×10^{-3}
AS-F7A	5.0×10^{10}	2.5×10^{9}	1550	6.2×10^{-5}
AS-Q8A	4.0×10^{10}	20×10^{9}	1400	7.0×10^{-5}
AS-Q9A	4.0×10^{10}	2.0×10^{9}	2500000	1.3×10^{-1}
AS-V10A	5.0×10^{10}	2.5×10^{9}	2250	9.0×10^{-5}
AS-C11S	2.0×10^{10}	1.0×10^{9}	200	2.0×10^{-5}
AS-C1SC11S*	7.0×10^{10}	3.5×10^{9}	4000	1.1×10^{-4}
C9-4	5.0×10^{10}	2.5×10^{9}	5250000	2.1×10^{-1}
Irrel.	4.0×10^{10}	2.0×10^{9}	150	7.5×10^{-6}

* AS-V3A and AS-C1SC11S were examined for SC binding in separate experiments with internal controls reaching the same yields as presented here.

4.6 Construction and expression of GFP fusion proteins with target binding capacity

An objective of this work has been to select short peptides with high affinity for SC that can mediate transcytosis. Thus, successful candidates have potential as fusion partners to therapeutically active molecules or diagnostics targeted to the mucosae. To reveal whether the characterized phage displayed peptide, C9-4, retain its SC binding capacity when taken out of the phage format, GFP fusion proteins were constructed.

Three, N-terminal, GFP fusion proteins were engineered alongside GFP without any fusion peptide. C9-4-Mono-GFP comprised one SC binding C9-4 peptide, selected by phage display, linked to GFP using a short spacer region. C9-4-Dimer-GFP was constructed as two C9-4 peptides connected by a linker and then coupled to GFP using a spacer region. The dimer was constructed to mimic the multimeric presentation of peptides seen on phage, consequently giving rise to increased avidity for SC. C1qC10-Mono-GFP contained a C1q binding peptide, selected by phage display, attached to GFP via a spacer region. The fusion free GFP is referred to as GFP only in the following. The GFP fusions were expressed in a prokaryotic system and screened for target binding.

4.6.1 Confirmation of fusion peptide sequences

Sequencing of the GFP-fusion peptide encoding inserts showed that GFP, C9-4-Mono-GFP and C1qC10-Mono-GFP matched the expected sequence. The C9-4-Dimer-GFP sequence, however, had two guanine (G) deletions, consequently leading to a frame shift mutation. The deletions were unsuccessfully attempted corrected. As a result the C9-4-Dimer-GFP was not expressed and excluded from further testing. The sequencing results are summarized in table 8. Detailed sequencing information is enclosed as Abi files in appendix 4.

Table 8. Confirmed sequences of peptides genetically fused to GFP

Fusion peptide	DNA sequence [a]
C9-4-Mono-	CCATGGGAGCCTGTGTGGTTTGGATGGGTTTTCAGCAGGTGTGTGGAGGACCAGGA-GFP
C9-4-Dimer-	CCATGGGAGCCTGTGTGGTTTGGATGGGTTTTCAGCAGGTGTGTXGAGGAGGACCAGGA [b] TGTXTGGTTTGGATGGGTTTTCAGCAGGTGTGTGGAGGACCAGGA-GFP
C1qC10-Mono-	CCATGGGAGCCTGTTATTGGGTGGGTACGTGGGGTGAGGCTGTTTGTGGAGGACCAGGA-GFP

a) The sequences are given in 5'-3' direction. Legend as in table 2.
b) The C9-4-Dimer was constructed with two C9-4 peptides genetically connected by a linker sequence (GGGPG). However, as indicated with X, the sequencing revealed a G deletion in the linker as well as in the repeated C9-4 monomer. Consequently a frame shift mutation arises.

4.6.2 Prokaryotic expression of GFP fusion proteins

C9-4-Mono-GFP, C1qC10-Mono-GFP and GFP, expressed in the prokaryotic pET system alongside GFP negative cells, were harvested three hours after induction and studied by fluorescence microscopy. C9-4-Mono-GFP, C1qC10-Mono-GFP and GFP cells were confirmed to express auto fluorescent GFP. Growth BL21DE expression host cells transformed with pET21d vector DNA without GFP encoding inserts (GFPneg) were confirmed by phase contrast microscopy. Microscopy photographs are presented in figure 7.

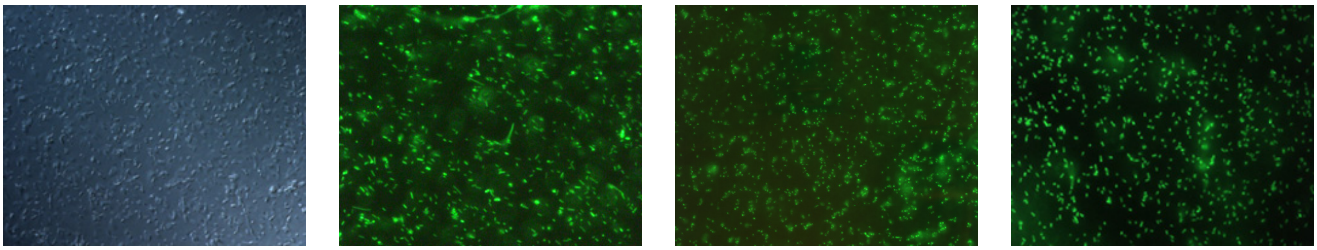


Figure 7. Detection of GFP expression by fluorescence microscopy
Cells harvested three hours after induction were screened for GFP expression using a fluorescence microscope (40x). A: GFPneg (phase contrast reveals actual cell growth). B: GFP. C: Cells expressing C9-4-Mono-GFP. D: Cells expressing C1qC10-Mono-GFP.

Additionally, to confirm the presence of a protein the size of GFP, harvested cells and cell lysates were further analysed by SDS PAGE. Commercially available, recombinant, GFP (kDa 27) was applied to the gel to estimate the amount of GFP in the cells and lysates respectively. SDS PAGE disclosed the presence of a protein the size of GFP (figure 8, line 2) and the somewhat larger proteins corresponding to C9-4-Mono-GFP (figure 8, line 3) and C1qC10-Mono-GFP (figure 8, line 4) respectively. Cells and lysate known to be GFP negative (figure 8, line 5) does not *Coomassie*^{TM} stain for proteins the size of GFP. The SDS PAGE results are shown in figure 8.

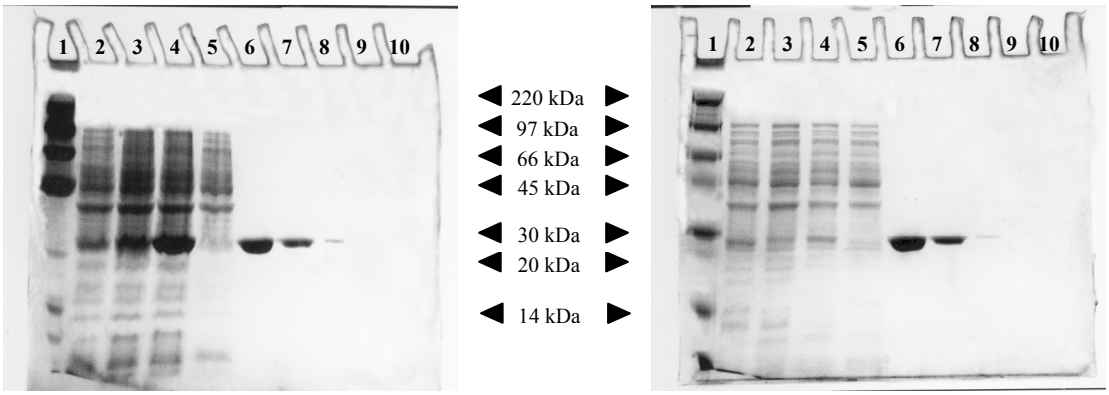


Figure 8. Expression of GFP fusion proteins
12% SDS PAGE gels prepared with resuspended cells (A) and lysates (B) respectively, were coomassie stained to verify expression of GFP and GFP fusions.

A/B1 MW
A/B2 GFP
A/B3 C9-4-Mono-GFP
A/B4 C1qC10-Mono-GFP
A/B5 GFPneg
A/B6 5µg rGFP (27kDa)
A/B7 1µg rGFP
A/B8 0.1µg rGFP
A/B9 0.01µg rGFP
A/B10 Empty.

Samples (12µL) of cells expressing GFP without peptide fusion (line A2) and cells expressing C9-4-Mono-GFP (line A3) holds approximately 1µg GFP (~80µg/mL), whereas the C1qC10-Mono-GFP (line A4) cell sample indicates approximately 5µg GFP (~400µg/mL). The concentration was determined by band intensity comparison with dilutions of commercial GFP (line A6-9). Samples (12µL) of lysates containing GFP without fusion peptide (line B2), C9-4-Mono-GFP (line B3) and C1qC10-Mono-GFP (line B4) all correspond roughly to 0.1µg commercially obtained GFP (8µg/mL) or a bit more. The expression level of C9-4-Mono-GFP appears to be somewhat lower than for the two other constructs. The GFP content of each individual lysate was determined as for the cells.

Finally, to confirm the presence of GFP, the lysates were investigated by western blotting using a GFP specific primary antibody. Dilutions of commercially obtained GFP were included as positive control and to estimate the amount of GFP in the lysates.

Western blotting confirmed the presence of GFP in the lysates from fusion free GFP (figure 9, line 2), C9-4-Mono-GFP (figure 9, line 3) and C1qC10-Mono-GFP (figure 9, line 4). Lysate from GFPneg cells gave no signal in the western blott (figure 9, line 6). The blotting results are presented in figure 9.

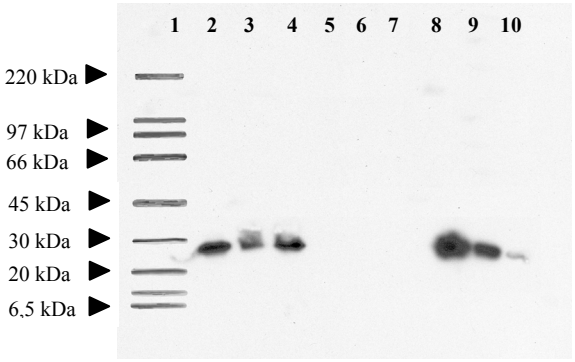


Figure 9. Detection of GFP in cell lysates by western blotting
Cell lysates, diluted 1:10 in PBS, were separated by SDS PAGE alongside dilutions of commercially available rGFP and transferred to a blotting membrane prior to detection using an anti GFP primary antibody.

1 MW	**6** GFPneg
2 GFP	**7** Empty
3 C9-4-Mono-GFP	**8** 10µg/mL rGFP
4 C1qC10-Mono-GFP	**9** 1µg/mL rGFP
5 Empty	**10** 0.1µg/mL rGFP

Note: The molecular weight ladder has been added digitally

12µL aliquots of GFP positive lysates, diluted 1:10 in PBS, correspond to the signal from 12µL of 1µg/mL commercial GFP (line 9). Thus the lysate concentration is calculated to ~10µg/mL GFP, concurring with the lysate concentration determined by *Coomassie*^{TM} staining. Western blotting confirms lower expression of C9-4-Mono-GFP compared to the two other constructs. In conclusion, C9-4-Mono-GFP, C1qC10-Mono-GFP and fusion free GFP cells all express auto fluorescent GFP.

4.6.3 Ammonium sulphate precipitation of GFP lysate

Lysate from BL21DE cells, expressing GFP without a fusion peptide, was ammonium sulphate precipitated to examine whether this method was a feasible purification procedure for the GFP lysates. Increasing amounts of ammonium sulphate were added to aliquots of lysate. Subsequently, the samples were examined by SDS PAGE and *Coomassie*^{TM} stained.

The precipitation study demonstrated that hardly any protein precipitate until 30% (1g/100mL =1%) of ammonium sulphate was added (figure 10, line 5). At this concentration however, essentially all proteins in the lysate precipitate. The results of the ammonium sulphate precipitation are presented in figure 10.

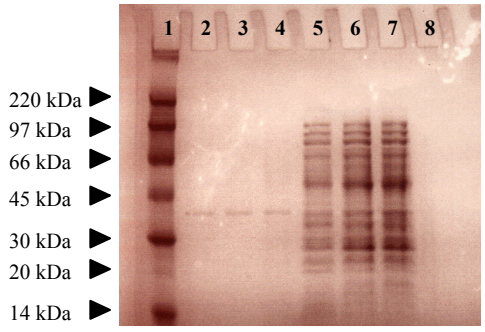


Figure 10. Ammonium sulphate precipitation

Lysate, containing GFP without fusion peptide (GFP), was Ammonium sulphate precipitated by adding increasing amounts of the salt (1g/100mL=1%). The precipitates were assayed by SDS PAGE and Coomassi staining.

1 MW	**5** 30%
2 5%	**6** 40%
3 10%	**7** 50%
4 20%	**8** Empty

Evidently, this system requires fine tuning to play a practical role in lysate purification. The reported non- or all precipitation renders the system ineffective for the intended application, thus it was not implemented for the other lysates.

4.7 Target binding capacity of GFP fusion proteins

To examine the fusion peptides' target binding capability when removed from the phage format in which they had been initially selected, C9-4-Mono-GFP and C1qC10-Mono-GFP were screened for their capacity to bind immobilized SC and C1q, respectively. C1qC10-Mono-GFP lysate was used as negative control in the SC biding studies whereas C9-4-Mono-GFP was used in the C1q binding assays. GFP lysate and GFPneg lysates were applied as controls in both binding assays.

4.7.1 Standardization of lysate fluorescence intensity

The concentration of GFP in the lysates was standardized by fluorescence measurement and subsequent dilution to ensure consistent GFP input levels. Fluorescence from undiluted C9-4-Mono-GFP lysate was measured to ~20000 RFU, corresponding to ~10µg/mL of commercially obtained GFP. Undiluted C1qC10-Mono-GFP lysate and GFP-lysate were measured to approximately 200000 RFU. According to RFU values from a dilution series, these lysates were diluted 1:7 and 1:12, respectively, to match the level of fluorescence measured for C9-4-Mono-GFP. The results of the lysate standardization are summarized in figure 11.

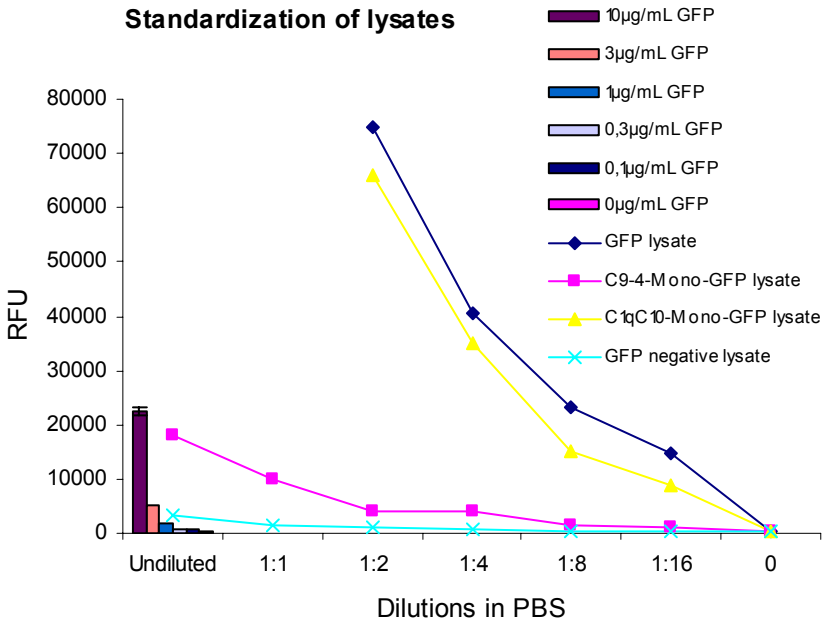


Figure 11. Standardization of lysates

Fluorescence in 100µL of dilutions of lysates (lines) and commercially available GFP (columns) were determined as relative fluorescence units (RFU) using a fluorometer. Undiluted C9-4-Mono-GFP lysate (pink line), ~20000RFU corresponds to 10µg/mL commercially available control GFP (purple column). Undiluted GFP (blue line) and C1qC10-Mono-GFP (yellow line) lysates were measured to ~200000RFU. 1:12 and 1:7 dilutions of GFP and C1qC10-Mono-GFP lysates, respectively, correspond to ~20000RFU (10µg/mL commercially obtained GFP). Background fluorescence from GFP negative lysate (GFPneg) is represented by the turquoise line.

Fluorescence measurement of the lysates revealed the presence of functional auto fluorescent GFP. Consistent with the SDS PAGE and Western blotting findings, C9-4-Mono-GFP contains less functional GFP than the two other GFP positive lysates. Lysates diluted to ~20000RFU, corresponding to approximately 10µg/mL of commercially obtained control GFP, were used as starting concentrations in target binding assays.

4.7.2 Binding to SC as determined by ELISA and fluorescence measurement

To determine whether C9-4 binds its target when shifted from the phage display context to the GFP fusion format, C9-4-Mono-GFP was screened for binding to immobilized SC using ELISA and fluorescence measurement. *Nunc-Immuno*^{TM} *MaxiSorp*^{TM} 96 well plates coated with 100µL of SC at 10µg/mL were used as solid phase in the ELISA. 100µL of each standardized lysate (~20000 RFU) were pre-incubated in 100µL SKM and added to two SC coated wells (10000 RFU/ reaction). Black *FluoroNunc*^{TM} *MaxiSorp*^{TM} 96 well plates coated with 100µL of SC at 10µg/mL were used as solid phase in the fluorescence measurement study. 100µL standardized lysate (20000 RFU) were added to each well in duplets.

No SC binding above background levels was detected in either of the assays. Accordingly, C9-4 does not appear to bind to SC in a detectable manner in this system when shifted from the phage display format to a monomeric GFP fusion. The results of C9-4-Mono-GFP binding to immobilized SC are presented in figure 12.

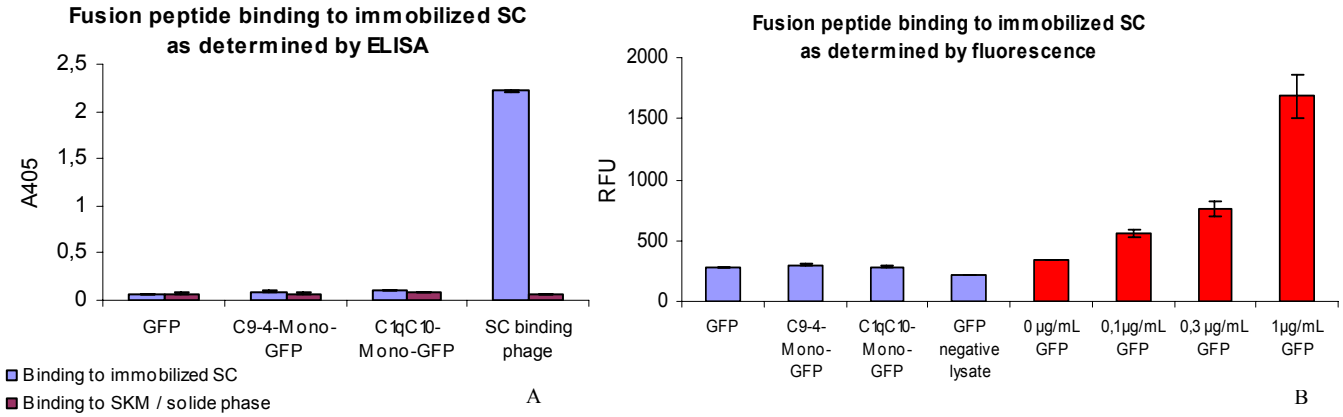


Figure 12. Fusion peptide binding to immobilized SC
A: Binding to immobilized SC was investigated by ELISA using an anti-GFP antibody. Standardized GFP positive lysates (~20000RFU/100µL) were preincubated (1:1) in blocking solution, thus 100µL input to each well correspond to ~10000RFU. The integrity of the coat was confirmed using a SC binding phage displaying C9-4.
SC binding and background binding is indicated in light blue and purple respectively.
B: Fusion peptide binding to SC (light blue) as determined by fluorescence. 100µL standardized GFP positive lysates and GFP negative lysate were incubated with immobilized SC. Following washing fluorescence was measured to reveal bound GFP fusion proteins. Dilutions of commercially obtained GFP (red) were included as positive control.

4.7.3 Binding to C1q as determined by ELISA and fluorescence measurement

A C1q binding peptide, selected by phage display, was fused to GFP to serve as a negative control in the SC binding studies. The peptide has previously been shown to bind C1q with high affinity when shifted from the phage display context to alternative fusion partners (Lunde et al., 2002). Consequently, we expected the C1qC10-Mono-GFP to bind its immobilized target. This construct would reveal whether target binding of GFP-fusion proteins, expressed bacterially, could be detected at all by fluorescence measurement in the concentrations we were obtaining.

To determine whether the C1q binding peptide, binds its target when presented in the GFP fusion format, C1qC10-Mono-GFP was screened for binding to immobilized C1q using ELISA and fluorescence measurement. *Nunc-Immuno*^{TM} *MaxiSorp*^{TM} 96 well plates coated with 100µL of C1q at 10µg/mL were used as solid phase in the ELISA. 100µL of each standardized lysate (~20000 RFU) were pre-incubated in 100µL SKM and added to two C1q coated wells (10000 RFU/ reaction). Black *FluoroNunc*^{TM} *MaxiSorp*^{TM} 96 well plates coated with 100µL of C1q at 10µg/mL were used as solid phase in the fluorescence measurement study. 100µL standardized lysate (20000 RFU) were added to each well in duplets.

Both the ELISA and fluorescence assay shows that C1qC10-Mono-GFP binds its target above background binding levels. Fluorescence readout suggests binding of GFP corresponding to approximately 0,3µg/mL of commercially obtained GFP. Thus, the phage display selected C1q binding peptide appears to bind its target also when displayed in an alternative fusion context. The results of the C1q binding assays are summarized in figure 13.

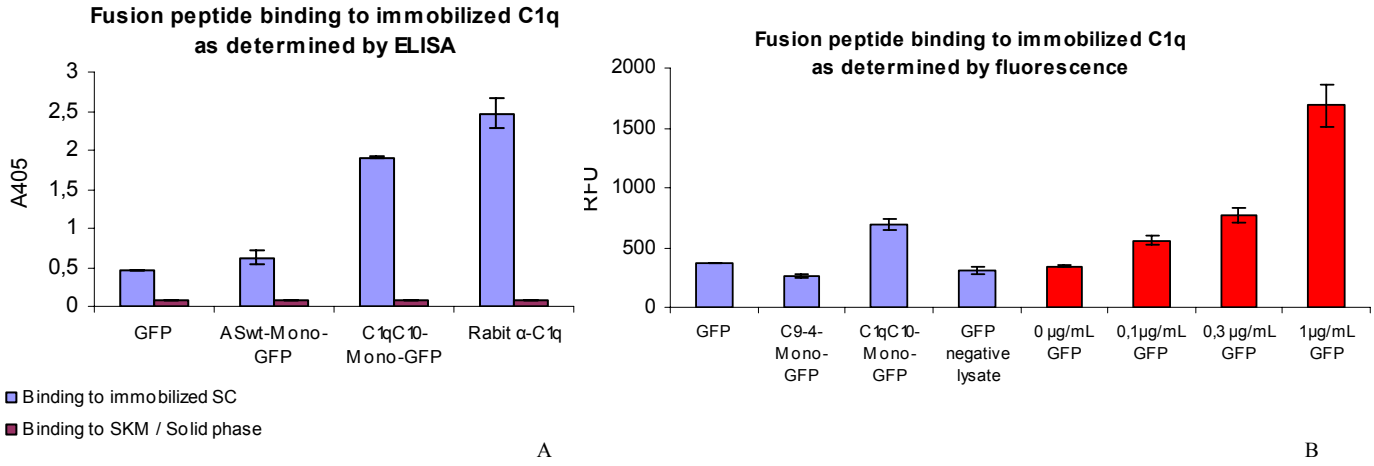


Figure 13. Fusion peptide binding to immobilized C1q
A: Binding to immobilized C1q (light blue) and background binding (purple) was investigated by ELISA using an anti-GFP antibody. Standardized GFP positive lysates (~20000RFU/100µL) were preincubated (1:1) in blocking solution, thus 100µL input to each well correspond to ~10000RFU. The integrity of the coat was confirmed using an anti-C1q antibody.
B: Fusion peptide binding to C1q was determined by fluorescence measurement. 100µL standardized GFP positive lysates and GFP negative lysate were incubated with immobilized C1q. Following washing fluorescence was measured to reveal bound GFP fusion proteins. Dilutions of commercially obtained GFP (red) were included as positive control.

In conclusion, the selected C1q binding control peptide binds its target and the interaction can be detected both by ELISA and by fluorescence measurement of bound GFP-fusion protein. Accordingly, it should be possible to detect target-fusion protein interactions using direct measurement of target bound auto-fluorescent fusion proteins, provided that the interaction is strong enough.

5. Discussion

Delivery of therapeutic macromolecules, gene therapy vectors and diagnostics to tissue may be hampered by endothelial or epithelial barriers. In an effort to overcome this obstacle, delivery strategies exploiting innate trans-barrier transport mechanisms have been suggested. To explore the possibility of selecting peptides targeting human pIgR a selection procedure for isolating human SC binding phages was developed and phages binding to SC were investigated for pIgR dependent transcytosis.

The isolated phage clone C9-4 was specifically transcytosed across MDCK cells expressing human pIgR. This is comparable to what has recently been reported by others on peptide targeted transport of intact phage particles across receptor expressing MDCK cells. A 9-mer phage peptide with an integrin binding motif has been demonstrated to yield 1000 to 10,000-fold improved transport compared to a control phage without the fusion peptide in basolateral to apical transport across MDCK cells (Ivanenkov and Menon, 2000). Likewise, peptides selected from a 20-mer and a 40-mer peptide phage display library by *in vitro* transcytosis across rabbit pIgR transfected MDCK cells, revealed increased transcytosis across receptor transfected cells compared to untransfected cells (White and Capra, 2002). One of the selected phages, displaying a peptide of 40 amino acids, was shown to be transported *in vivo* in a rat model measuring hepatic bile transport (White and Capra, 2002). Furthermore, GFP fusion proteins comprising peptides selected for pIgR binding from a IgA gene fragment library (Hexham et al., 1999) were shown to be specifically transcytosed across pIgR transfected MDCK cells (White and Capra, 2002).

To identify short novel peptides with capacity to target human pIgR, a 7-mer and Cys9 phage display library were subjected to two rounds of affinity selection and single clone screening for binding to human SC purified from colostrum. For the 7-mer library a SC specific 5-fold increase in recovery from the first round of selection to the second selection round was seen when using SKM as blocking agent, indicating enrichment in the phage pool for SC binding clones. However, as an ELISA assay failed to confirm the presence of SC binding phages in the amplified second eluates, no attempt were made to screen for individual target binding clones.

Considering that only ~20 copies of each phage were added to the first round of selection, it would be of interest to investigate whether consecutive rounds of selection and / or altered selection conditions (i.e. stringency) would have given a clearer trend of enrichment.

Regarding the Cys9 library, approximately 280 copies of each phage clone were present in the initial round of affinity selection. The peptides are expected to fold into a more rigid conformation due to disulfide bridge formation. This has proven to be advantageous when screening for high affinity target binding peptides (McLafferty et al., 1993, review; O'Neil et al., 1992). Despite the fact that no enrichment was indicated by the recovery rate after two rounds of selection, an ELISA assay revealed the presence of detectable amounts of SC binding phage in the amplified Cys9 second round eluates. When using BSA as blocking agent, significant background binding was observed indicating enrichment of BSA or plastic binding phage particles. No such background binding was observed for eluates derived from selections using SKM as blocking reagent. Cross-blocking with SKM demonstrated that the amplified Cys9 eluates, derived from selections using BSA as blocking agent, in fact contain some specific SC binding phages. Although, SKM reduced the signal for specific target binding, SKM seemed to be the better choice of blocking agent with regard to recovery of specific binders.

Automation of phage display library selection and high throughput screening has proven useful in the study of receptor ligand interactions and has resulted in the identification of a broad array of novel peptides and engineered proteins (Hallborn and Carlsson, 2002, review; Krebs et al., 2001; Sidhu et al., 2003). Single clones were successfully propagated in high throughput compatible deep well growth plates. Increased evaporation from the outermost wells was observed, but all wells appeared to produce cultures of approximately equal density. A BioWorks script (Appendix 1) designed to direct the Biomek2000 robot to aliquot phage supernatants was tested and could be further developed to be included in a high throughput phage display screening protocol.

352 single clones were screened for binding to immobilized SC by ELISA. Clones, positive in the initial high throughput screening were re-screened to confirm specific target binding. The confirmative ELISA revealed that all, but two, of the 43 clones selected for re-screening bound specifically to SC rather than the background. 7 different sequences were identified amongst the SC specific clones.

Only two of these, CIIVPHAYAWC and CVVWMGFQQVC, encoded 9-mer peptides flanked by cysteins. CIIVPHAYAWC was recovered six times from individual clones originating from the second round of selection using SKM as blocking reagent. CVVWMGFQQVC, isolated 13 times from separate clones, was found amongst positive clones after both the first and second round of selection using SKM as blocking agent. Interestingly, this dominating sequence turned out to be identical to clone C9-4, previously identified by Braathen et al., (unpublished data) using the same Cys9 library. C9-4 was used as a positive control phage in the screening procedures. Thus, it can not be excluded that its re-occurrence may be due to contamination. However, the re-isolation may also indicate that the sequence is one of the better SC binders in the library, if not the best, thus it should be expected to dominate.

A 14-mer peptide CALVSEAGCLVWAA named ASR, deviating from the expected cystein constrained 9-mer sequence, was identified 20 times amongst phages originating from all amplified eluates. All of the three described sequences contain two cysteins. The four different 7-mer peptides RFWWGWY, PFVVLLV, RGPFVYM and GWAGWLG each isolated only once are likely to be contaminants originating from the 7-mer library used in parallel to the Cys9 library. Two selected clones were negative for SC binding in the confirmative SC binding ELISA; these were identified to be the latter two of the 7-mer contaminants.

The Cys9 library (Lauvrak et al., 2003) has previously been the source of a number of different peptides with clear consensus motifs covering a stretch of four to six amino acids when using monoclonal antibodies as target in selection (Lauvrak et al., 2003). However, when using SC (this work and Braathen et al., unpublished) and FcγRI (Berntzen et.al, unpublished) as target no consensus motifs were found among the target binding clones. The library size was estimated to 5×10^7 different clones, which represents only a fraction of all theoretically possible 9-mer peptides. A Cys6 library of the same size gave consensus motifs involving more than five different peptides and five out of the six amino acids for peptides binding to FcγRI and FcγRIIa (Berntzen et al., unpublished). No SC binding sequence with consensus motifs of only six amino acids could be isolated from the same Cys6 library (Braathen et al., unpublished). Instead longer peptides, assumingly representing duplicates of inserts, were identified to have SC binding ability. Thus, it seems likely that a SC binding consensus motif would involve stretches of more than four to six amino acids. Notably, no consensus was observed amongst the 20-mer and 40-mer peptides isolated by White and Capra (2002) when targeting rabbit pIgR.

Approximately 10^9 different clones would be necessary to cover a complete presentation of all possible 6-mer sequences involving a NNK codon, as used for the libraries in this work. Considering the fact that random libraries can be constructed with diversities exceeding 10^9 different sequences (Noren and Noren, 2001), our Cys9 library is relatively small. This may potentially explain why not additional SC binding peptides were identified. Screening of larger libraries might yield additional SC binding peptides with consensus motifs.

Notably, IgA binds covalently to SC involving cysteines (Weicker and Underdown, 1975). The presence of cysteins in SC and the library may potentially lead to covalent disulfide interaction during the affinity selection procedure. Thus, potentially strong SC binders may never have been recovered by acidic elution. The presence of such peptides has not been confirmed. However, infection by adding fresh host cells directly to the selection wells, as an alternative to acidic elution, might lead to the identification of such SC binding candidates.

Alternative strategies to obtain SC binding peptides may include randomization between rounds of affinity selection. This process mimics the affinity maturation seen for antibodies (Hawkins et al., 1992). Based on the observation that no short, consensus motif containing peptides were not selected, the binding sites on SC might be very different from the hyper variable regions of an antibody with its protruding CDR loops. Thus, to increase affinity the surface area available for receptor interaction can be expanded by adding randomized NNK codons both N- and /or C-terminally to a promising peptide candidate. Another approach could be the use of cells transfected with pIgR as solid phase in the affinity selection. Selection directly on cells allows for a native presentation of the receptor, thus additional binding sites may occur. Subsequent labeling for bound phage particles, cell sorting and host cell infection may lead to the identification of additional pIgR binding peptides not identified by means of classical selection in plastic wells. Yet another potential method is the selection of pIgR binding peptides, and screening of such, by transcytosis across pIgR transfected polarized cells. This allows for direct recovery of receptor binding phage from the apical side of the cell growth chamber. Selection by transcytosis would be highly relevant in our work, as our aim is to select peptides that are transported across epithelia. This method was successfully used by White and Capra (2002) to obtain SC binding peptides, 20 and 40 amino acids of length, from random libraries.

The SC binding capacity of the two most dominating phages, in terms of number of identical clones isolated, C9-4 (CVVWMGFQQVC) and ASR (CALVSEAGCLVWAA), were investigated in a target binding inhibition assay (figure 4). The dIgA and pIgM interaction with SC has been characterized in detail and the N-terminal domain 1 of pIgR has been demonstrated to be the dIgA and pIgM binding domain (Bakos et al., 1994). Furthermore, amino acids 402-410 (QEPSQGTTT) of domain Cα3 of IgA has been suggested to be critical for pIgR binding (Hexham et al., 1999). In addition, the J-chain has been confirmed to be of importance in the binding of pIgs to their receptor and thus formation of secretory Abs (Johansen et al., 2001). Accordingly, to map whether C9-4 and ASR interact with SC at a binding site utilized by pIgs, phage binding was attempted hampered by addition of dIgA and pIgM. Binding of C9-4 or ASR to immobilized SC is not affected by polymeric Igs, indicating that the selected phages do not bind to SC at the same site or at overlapping sites as those utilized by pIgM or sIgA.

The virulence factor SpsA of *S. pneumonia*, has been demonstrated to bind sIgA and fSC (Hammerschmidt et al., 1997). The interaction has been demonstrated to be specific for the human pIgR (Hammerschmidt et al., 2000) and recent work has demonstrated that domains D3 and D4 of human pIgR together are necessary and sufficient for SpsA binding (Elm et al., accepted for publication; Lu et al., 2003). SpsA hexapeptide (YRNYPT), located between amino acids 198 and 203, and in particular the highly conserved YPT motif, has been demonstrated to be of critical importance for pIgR binding. However, the recent results from Lu et al., strongly suggests that the SpsA binding motif(s) on pIgR is of conformational nature, indicating the involvement of more than the YPT motif in the receptor-ligand interaction. Three fragments of SpsA were applied as inhibitors. Approximately 10 nM of the SpsA fragments SH2 and SM1 reduced SC binding by 50% (A_{405} value reduction), whereas the SC negative SpsA fragment, SH3, did not interfere with the binding, suggesting that C9-4 binds SC in a SpsA like manner. C9-4 does not contain an YPT motif, indicating that the binding site utilized by SpsA and C9-4 is not identical. However, the observation that C9-4 binding to SC clearly is inhibited by SpsA suggests that the binding sites overlap to some extent or potentially that occupation of one of the bindings sites renders the other unavailable due to sterical hindrance or conformational change. C9-4 binding to series of human/murine chimeras of pIgR may further elucidate how and where C9-4 binds to SC (work in progress). ASR binding to SC was not affected by any of the applied inhibitors, indicating that the phage binds SC at a site not utilized by polymeric Igs or by SpsA.

C9-4 and ASR's ability to bind SC, when presented in a more native configuration basolaterally on MDCK cells, and their ability to be transcytosed across this in vitro epithelial model was examined to asses their prospects for use in mucosal targeting. To reveal whether the phage-receptor interaction, like the SpsA-pIgR interaction, is species specific both murine and human SC were included in the transcytosis experiments. We demonstrate that C9-4 is specifically transcytosed across MDCK cells transfected with the human pIgR only. Accordingly, C9-4 interacts with SC in a species dependent manner despite considerable amino acid homology (55%) between the human and murine receptor (Hammerschmidt et al., 2000; Zhang et al., 2000), adding further similarities to the list of traits shared with SpsA. The fact that untransfected MDCK cells are unable to mediate C9-4 transcytosis strongly suggests that the pIgR mediated transcytosis is the actual transport mechanism in action. No receptor specific transfer of ASR across polarized MDCK cells was seen. Whether this is due to lack of SC binding in the cellular assay or is caused by some other unknown limitation has not been examined.

The contribution of each individual amino acid on peptide binding to SC was examined by substitution mutation analysis. Flanking cysteins were substituted with serine whereas the remaining amino acids were systematically substituted with alanine. Alanine substitution does not interfere with the protein backbone, making it a suitable tool for studying receptor-ligand interaction, protein function and protein stability (reviewed in Morrison and Weiss, 2001). Serine substitution of cysteine implicates the loss of ability to form disulfide bridges, thus the mutants are expected to lack a rigid conformation, potentially of importance for binding capacity. All mutants, but three, completely lost their ability to bind SC in an ELISA detectable manner. Mutant AS-V3A provided an A_{405} readout signal approximately 40% of the signal seen for the positive control, C9-4. Furthermore, mutant AS-G6A binding to SC gave a signal approximately 70% of the signal seen for the positive control and AS-Q9A gave a SC binding signal identical to that of C9-4. Elution-titration, performed in parallel to the ELISA, revealed the same tendency. However, this assay revealed that all C9-4 amino acid substitution mutants did bind SC slightly better than an irrelevant phage. Furthermore a difference between C9-4 and AS-Q9A not observed with the ELISA assay was uncovered, as recovery for AS-Q9A was only 60% of the yield seen for C9-4 wild type phage. A larger difference in recovery rates for AS-V3A and AS-G6A was seen than would have been expected from the ELISA results. Only 1.5% and 2% of the recovery seen for C9-4 was observed for AS-V3A and AS-G6A, respectively.

Thus, the elution-titration study has higher resolution than the ELISA assay. The reported variation in ELISA signal and phage recovery can not be accounted for by considering a slight difference of amount of input phage only, thus the variation is likely to be due to the functionality of the peptides. Alanine for valine substitution as in AS-V3A implicates the substitution of $CH(CH_3)_2$ for a CH_3 side chain. These are both nonpolar, hydrophobic side chains and both amino acids are likely to cluster together on the inside of a structurally constrained polypeptide. Consequently, the reported substitution was likely not to affect the overall structure of the peptide. Notably, a substitution of V in position 2 and 9 was not tolerated, indicating that the more hydrophobic nature of V relative to A at these positions might be of more importance than in position 3. Alanine for glycine substitution implies the removal of a CH_3 group in favour of a simple H atom. Glycine lacks a functional side chain group and is not expected to energetically contribute to the interaction, but might be of great importance for the structure. Substitution of the uncharged, polar side chain of glutamine $(CH)_2CONH_2$ with a non-polar CH_3 group implicates the removal of a charged protruding, slightly hydrophilic amino acid in exchange for a small hydrophobic one. Consequently, when such a mutation does not affect the peptides ability to bind SC in a serious manner, the amino acid can not be directly involved in the interaction. Mutants AS-C1S, AS-C11S and the double mutant AS-C1SC11S did not bind SC, suggests an important role of disulphide bridge formation on C9-4 binding to SC. The apparent requirement for cysteines suggests that C9-4 folds into a restrained tertiary structure and that this cysteine constrained conformation is essential for target binding.

A synthetic version of the gpIII presented C9-4 peptide inhibits binding of C9-4 phage to immobilized SC. 20nM (0.3µg/mL) of synthetic C9-4 reduced phage binding to SC by 50%, indicating that the peptide binds its target with remarkably high affinity (Braathen et al., unpublished). It should be noted that the nature of the inhibitor, in terms of whether it is present as monomeric peptide or multimeric aggregate, has not been established. Thus, the results must be interpreted with this in mind.

To investigate whether C9-4 retained its SC binding capacity when shifted from the multimeric gpIII phage fusion format to a different fusion context, N-terminal GFP fusions were constructed and the fusion proteins ability to bind SC was examined. We aimed at constructing a monomeric fusion protein and a C9-4 dimer-GFP fusion. The dimer was included as a dimerization might potentially cause increased avidity for the target.

GFP fusions are frequently constructed with a spacer sequence between the auto fluorescent protein and its fusion partner (Cubitt et al., 1995). We opted for a flexible spacer with ability to form kinks to make sure C9-4 was protruding away from its much bigger fusion partner (238aa), thus a spacer based on the GGGGGGPG-sequence used by White and Capra (2002) was chosen. The two C9-4 peptides making up the dimmer fusion protein were inter-connected by a linker sequence almost identical to the applied spacer sequence. The choice of linker and spacer length was limited by the chosen cloning strategy. To limit the length of the synthetic oligonucleotides, a GGPG sequence was inserted between GFP and C9-4 whereas GGGPG was inserted between the two C9-4 monomers. We did not succeed in producing a C9-4 dimer fusion to GFP. The C9-4-Dimer-GFP fusion construct was confirmed to carry two deletion mutations, consequently leading to a frame shift mutation. Although we initially feared the formation of secondary structures on the very long mutagenic oligonucleotides, such a structure can not explain the deletion of two single nucleotides separated by 17 perfectly incorporated nucleotides. Alternative mutagenesis in one or more steps would most certainly correct the reported errors. However, such an approach was not attempted and the construction of a C9-4 dimer fusion was put to an end.

Crude lysates containing C9-4-Mono-GFP and control GFP fusion proteins was successfully obtained (figures 7, 8 and 9). In an effort to purify the GFP containing lysates, the use of ammonium sulphate precipitation was examined. At 30% *E.coli* proteins started to precipitate. Increasingly higher Coomassie staining intensity was seen for lysate aliquots added 40% and 50% of ammonium sulphate, respectively. Although no specific precipitation of the 27kDa sized GFP was observed, it may be possible to fine tune the precipitation conditions in order to obtain a GFP enriched fraction to be further purified by other means such as gel filtration. Alternatively, other approaches, such as heat denaturation and removal of coagulated *E.coli* proteins, can be tried as GFP is reported to tolerate heating to 78^{o}C before a 50% loss of fluorescence is seen (Tsien, 1998, review). Thus, *E.coli* proteins that denatures and coagulates at temperatures lower than that of GFP might be removed from the cell lysates by centrifugation. Yet another possibility lies in the unique structure of GFP that allows for isolation of the active fluorescent protein directly from the crude cellular sources by extraction with organic solvents (Skosyrev et al., 2003). White and Capra (2002), reported to use crude cell lystate with GFP fusion proteins in their pIgR targeting study.

To estimate the amount of functional GFP in the lysates, dilutions of fusion protein lysate was compared with dilutions of commercially obtained GFP of known concentration. The concentration determination was hampered by the prize of GFP. Consequently an insufficient amount of GFP with known concentration was obtained. As a result the standard dilution failed to reach a linear phase. As C9-4-Mono-GFP reached the lowest fluorescence intensity level (~20,000 RFU), corresponding to approximately 10µg/mL of commercial GFP), control lysates were adjusted accordingly (figure 11).

We were unable to detect C9-4-Mono-GFP binding to immobilized SC by ELISA or direct fluorescence measurement (figure 12). It is reasonable to speculate that the lack of binding may be caused by too low affinity for the target. It should be noted that GFP tend to form dimers spontaneously (Tsien, 1998, review). The contribution of this property on multivalent peptide display has not been studied, neither has the potential presence of GFP dimers in the lysates been examined. The lack of target binding may also be caused by improper peptide folding or accessibility. The integrity of the fusion peptide was not inspected, thus it is not known whether the C9-4 peptide is actually present on the larger GFP molecule that we screened for. This could have been analyzed by N-terminal protein sequencing, which would have required a purification step. Yet another concern is the concentration of functional GFP-fusion peptide. The highest available concentration of the fusion protein was applied, thus purification and up concentration would potentially reveal the answer to this question.

A C1q binding peptide, identified by phage display (Bremnes et al., 1998), was applied as a negative control in the SC binding experiments. In elution-titration studies, phage particles displaying the control peptide were shown to yield recovery rates for C1q binding approximately 10-fold higher than what we observed for C9-4 (results not shown). Thus the avidity of the peptide on phage for its target seems to be higher for this peptide than observed for C9-4 and SC. Moreover, this peptide had previously been demonstrated to retain C1q binding capacity as Fab- and scFv fusions (Lunde et al., 2002). Accordingly, the chosen peptide was selected to serve as an overall system control to reveal the possibility of shifting peptides to GFP and still see target binding. As we successfully demonstrated that the control GFP fusion, C1qC10-Mono-GFP, did bind immobilized C1q in a manner detectable by both ELISA and fluorescence measurement the system worked. Thus the lack of C9-4 binding to SC could be due to low affinity, low availability or both.

A synthetic version of C9-4 conjugated to biotin was made. This version of the peptide did not dissolve in water (not shown). Solubilisation at high concentration in DMSO and subsequent dilutions in an adequate buffer might solve the problem. Specific transcytosis of the biotinylated peptide, or multimers thereof (e.g. aggregation by streptavidin binding), across pIgR transfected MDCK cells might reveal whether the identified sequence has the potential to target mucosa when removed from the phage display context.

In conclusion: Five phage display peptides binding to human SC were isolated. One of these peptides was shown to specifically bind human SC in a SpsA like manner. Phages displaying this peptide are specifically transcytosed across MDCK cells expressing human pIgR. This demonstrates that the scenario of using phage display derived peptides in mucosal targeting is achievable. However, as a GFP fusion of the peptide did not lead to detectable SC binding, further work needs to be done to identify a general pIgR targeting peptide.

6. References

Bakos M. A., Widen S. G. and Goldblum R. M. (1994) Expression and purification of biologically active domain I of the human polymeric immunoglobulin receptor. *Mol Immunol* **31**, 165-8.

Breitfeld P. P., Harris J. M. and Mostov K. E. (1989) Postendocytotic sorting of the ligand for the polymeric immunoglobulin receptor in Madin-Darby canine kidney cells. *J Cell Biol* **109**, 475-86.

Bremnes T., Lauvrak V., Lindqvist B. and Bakke O. (1998) Selection of phage displayed peptides from a random 10-mer library recognising a peptide target. *Immunotechnology* **4**, 21-8.

Braathen R., Sorensen V., Brandtzaeg P., Sandlie I. and Johansen F. E. (2002) The carboxyl-terminal domains of IgA and IgM direct isotype-specific polymerization and interaction with the polymeric immunoglobulin receptor. *J Biol Chem* **277**, 42755-62.

Cubitt A. B., Heim R., Adams S. R., Boyd A. E., Gross L. A. and Tsien R. Y. (1995) Understanding, improving and using green fluorescent proteins. *Trends Biochem Sci* **20**, 448-55.

Deperthes D. (2002) Phage display substrate: a blind method for determining protease specificity. *Biol Chem.* **383**, 1107-12.

Elm C., Braathen R., Bergmann S., Frank R., Vaerman J. P., Kaetzel C. S., Chhatwal G. S., Johansen F. E. and Hammerschmidt S. (accepted for publication) Ectodomains 3 and 4 of human polymeric immunoglobulin receptor (hpIgR) mediate invasion of Streptococcus pneumoniae into the epithelium. *J. Biol. Chem*.

Hallborn J. and Carlsson R. (2002) Automated screening procedure for high-throughput generation of antibody fragments. *Biotechniques* **Suppl**, 30-7.

Hammerschmidt S., Talay S. R., Brandtzaeg P. and Chhatwal G. S. (1997) SpsA, a novel pneumococcal surface protein with specific binding to secretory immunoglobulin A and secretory component. *Mol Microbiol* **25**, 1113-24.

Hammerschmidt S., Tillig M. P., Wolff S., Vaerman J. P. and Chhatwal G. S. (2000) Species-specific binding of human secretory component to SpsA protein of Streptococcus pneumoniae via a hexapeptide motif. *Mol Microbiol* **36**, 726-36.

Hawkins R. E., Russell S. J. and Winter G. (1992) Selection of phage antibodies by binding affinity. Mimicking affinity maturation. *J Mol Biol* **226**, 889-96.

Hetian L., Ping A., Shumei S., Xiaoying L., Luowen H., Jian W., Lin M., Meisheng L., Junshan Y. and Chengchao S. (2002) A novel peptide isolated from a phage display library inhibits tumor growth and metastasis by blocking the binding of vascular endothelial growth factor to its kinase domain receptor. *J Biol Chem* **277**, 43137-42.

Hexham J. M., White K. D., Carayannopoulos L. N., Mandecki W., Brisette R., Yang Y. S. and Capra J. D. (1999) A human immunoglobulin (Ig)A calpha3 domain motif directs polymeric Ig receptor-mediated secretion. *J Exp Med* **189**, 747-52.

Ihle O. and Michaelsen T. E. (2000) Efficient purification of DNA fragments using a protein binding membrane. *Nucleic Acids Res* **28**, E76.

Ivanenkov V. V. and Menon A. G. (2000) Peptide-mediated transcytosis of phage display vectors in MDCK cells. *Biochem Biophys Res Commun* **276**, 251-7.

Johansen F. E., Braathen R. and Brandtzaeg P. (2000) Role of J chain in secretory immunoglobulin formation. *Scand J Immunol* **52**, 240-8.

Johansen F. E., Braathen R. and Brandtzaeg P. (2001) The J chain is essential for polymeric Ig receptor-mediated epithelial transport of IgA. *J Immunol* **167**, 5185-92.

Kaetzel C. S. (2001) Polymeric Ig receptor: defender of the fort or Trojan horse? *Curr Biol* **11**, R35-8.

Kipriyanov S. M., Moldenhauer G. and Little M. (1997) High level production of soluble single chain antibodies in small-scale Escherichia coli cultures. *J Immunol Methods* **200**, 69-77.

Krebs B., Rauchenberger R., Reiffert S., Rothe C., Tesar M., Thomassen E., Cao M., Dreier T., Fischer D., Hoss A., Inge L., Knappik A., Marget M., Pack P., Meng X. Q., Schier R., Sohlemann P., Winter J., Wolle J. and Kretzschmar T. (2001) High-throughput generation and engineering of recombinant human antibodies. *J Immunol Methods* **254**, 67-84.

Lauvrak V., Berntzen G., Heggelund U., Herstad T. K., Sandin R. H., Dalseg R., Rosenquist E., Sandlie I. and Michaelsen T. J. (2003) Selection and characteriazation of cyclic peptides that bind to a monoclonal antibody against meningococcal L3,7,9 LPS. *Scand J Immunol*, Accepted for publication.

Lesinski G. B. and Westerink M. A. (2001) Novel vaccine strategies to T-independent antigens. *J Microbiol Methods.* **47**, 135-49.

Livnah O., Stura E. A., Johnson D. L., Middleton S. A., Mulcahy L. S., Wrighton N. C., Dower W. J., Jolliffe L. K. and Wilson I. A. (1996) Functional mimicry of a protein hormone by a peptide agonist: the EPO receptor complex at 2.8 A. *Science* **273**, 464-71.

Lu L., Lamm M. E., Li H., Corthesy B. and Zhang J. R. (2003) The human polymeric immunoglobulin receptor binds to Streptococcus pneumoniae via domains 3 and 4. *J Biol Chem* **278**, 48178-87.

Lunde E., Lauvrak V., Rasmussen I. B., Schjetne K. W., Thompson K. M., Michaelsen T. E., Brekke O. H., Sollid L. M., Bogen B. and Sandlie I. (2002) Troybodies and pepbodies. *Biochem Soc Trans* **30**, 500-6.

Luton F. and Mostov K. E. (1999) Transduction of basolateral-to-apical signals across epithelial cells: ligand-stimulated transcytosis of the polymeric immunoglobulin receptor requires two signals. *Mol Biol Cell* **10**, 1409-27.

McCarthy C. Chromas. Freely distributed software, South Port, Queensland.

McLafferty M. A., Kent R. B., Ladner R. C. and Markland W. (1993) M13 bacteriophage displaying disulfide-constrained microproteins. *Gene* **128**, 29-36.

Meissner P. S., Sisk W. P. and Berman M. L. (1987) Bacteriophage lambda cloning system for the construction of directional cDNA libraries. *Proc Natl Acad Sci U S A* **84**, 4171-5.

Morrison K. L. and Weiss G. A. (2001) Combinatorial alanine-scanning. *Curr Opin Chem Biol* **5**, 302-7.

Natvig I. B., Johansen F. E., Nordeng T. W., Haraldsen G. and Brandtzaeg P. (1997) Mechanism for enhanced external transfer of dimeric IgA over pentameric IgM: studies of diffusion, binding to the human polymeric Ig receptor, and epithelial transcytosis. *J Immunol* **159**, 4330-40.

Nilsson F., Tarli L., Viti F. and Neri D. (2000) The use of phage display for the development of tumour targeting agents. *Adv Drug Deliv Rev.* **43**, 165-96.

Norderhaug I. N., Johansen F. E., Schjerven H. and Brandtzaeg P. (1999) Regulation of the formation and external transport of secretory immunoglobulins. *Crit Rev Immunol* **19**, 481-508.

Noren K. A. and Noren C. J. (2001) Construction of high-complexity combinatorial phage display peptide libraries. *Methods* **23**, 169-78.

O'Neil K. T., Hoess R. H., Jackson S. A., Ramachandran N. S., Mousa S. A. and DeGrado W. F. (1992) Identification of novel peptide antagonists for GPIIb/IIIa from a conformationally constrained phage peptide library. *Proteins* **14**, 509-15.

Petrenko V. A. and Vodyanoy V. J. (2003) Phage display for detection of biological threat agents. *J Microbiol Methods.* **53**, 253-62.

Phalipon A. and Corthesy B. (2003) Novel functions of the polymeric Ig receptor: well beyond transport of immunoglobulins. *Trends Immunol* **24**, 55-8.

Roe M., Norderhaug I. N., Brandtzaeg P. and Johansen F. E. (1999) Fine specificity of ligand-binding domain 1 in the polymeric Ig receptor: importance of the CDR2-containing region for IgM interaction. *J Immunol* **162**, 6046-52.

Scott J. K. and Smith G. P. (1990) Searching for peptide ligands with an epitope library. *Science* **249**, 386-90.

Sidhu S. S., Fairbrother W. J. and Deshayes K. (2003) Exploring protein-protein interactions with phage display. *Chembiochem* **4**, 14-25.

Skosyrev V. S., Rudenko N. V., Yakhnin A. V., Zagranichny V. E., Popova L. I., Zakharov M. V., Gorokhovatsky A. Y. and Vinokurov L. M. (2003) EGFP as a fusion partner for the expression and organic extraction of small polypeptides. *Protein Expr Purif* **27**, 55-62.

Smith G. P. (1994) fd-tet: parent of all the vectors. Personal communication.

Smith G. P. and Scott J. K. (1993) Libraries of peptides and proteins displayed on filamentous phage. *Methods Enzymol* **217**, 228-57.

Tsien R. Y. (1998) The green fluorescent protein. *Annu Rev Biochem* **67**, 509-44.

Weicker J. and Underdown B. J. (1975) A study of the association of human secretory component with IgA and IgM proteins. *J Immunol* **114**, 1337-44.

White K. D. and Capra J. D. (2002) Targeting mucosal sites by polymeric immunoglobulin receptor-directed peptides. *J Exp Med* **196**, 551-5.

Zhang J. R., Mostov K. E., Lamm M. E., Nanno M., Shimida S., Ohwaki M. and Tuomanen E. (2000) The polymeric immunoglobulin receptor translocates pneumococci across human nasopharyngeal epithelial cells. *Cell* **102**, 827-37.

Zurita A. J., Arap W. and Pasqualini R. (2003) Mapping tumor vascular diversity by screening phage display libraries. *J Control Release.* **91**, 183-6.

7. Appendix

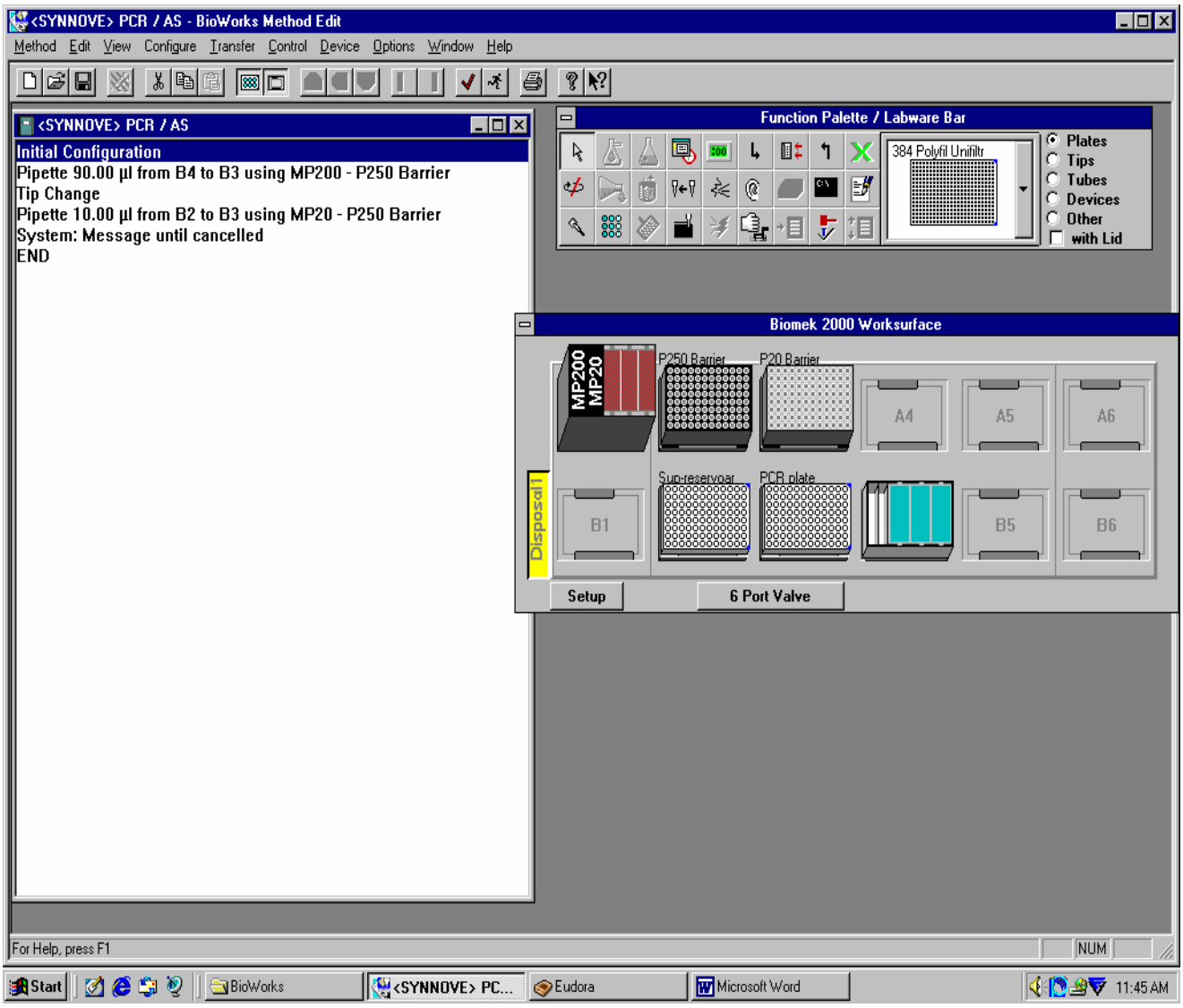
- **Appendix 1**
  BioWorks scripts for Biomek 2000 liquid handler designed for semi-automated sample handling and enhanced throughput.
- **Appendix 2**
  Abi files, presenting the sequences of peptides recovered from Cys9 affinity selection for SC binders.
- **Appendix 3**
  Abi files, confirming the systematic alanine- and serine substitution of amino acids of peptide C9-4.
- **Appendix 4**
  Abi files, presenting the sequences of the constructed GFP fusion peptides

Appendix 1

Three scripts, instructing Biomek 2000 (Beckman Coulter) to set up a PCR reaction, to aliqoute phage supernatnat and to run an ELISA, repectively, were designed for semi-automated sample handling and incrased throughput.

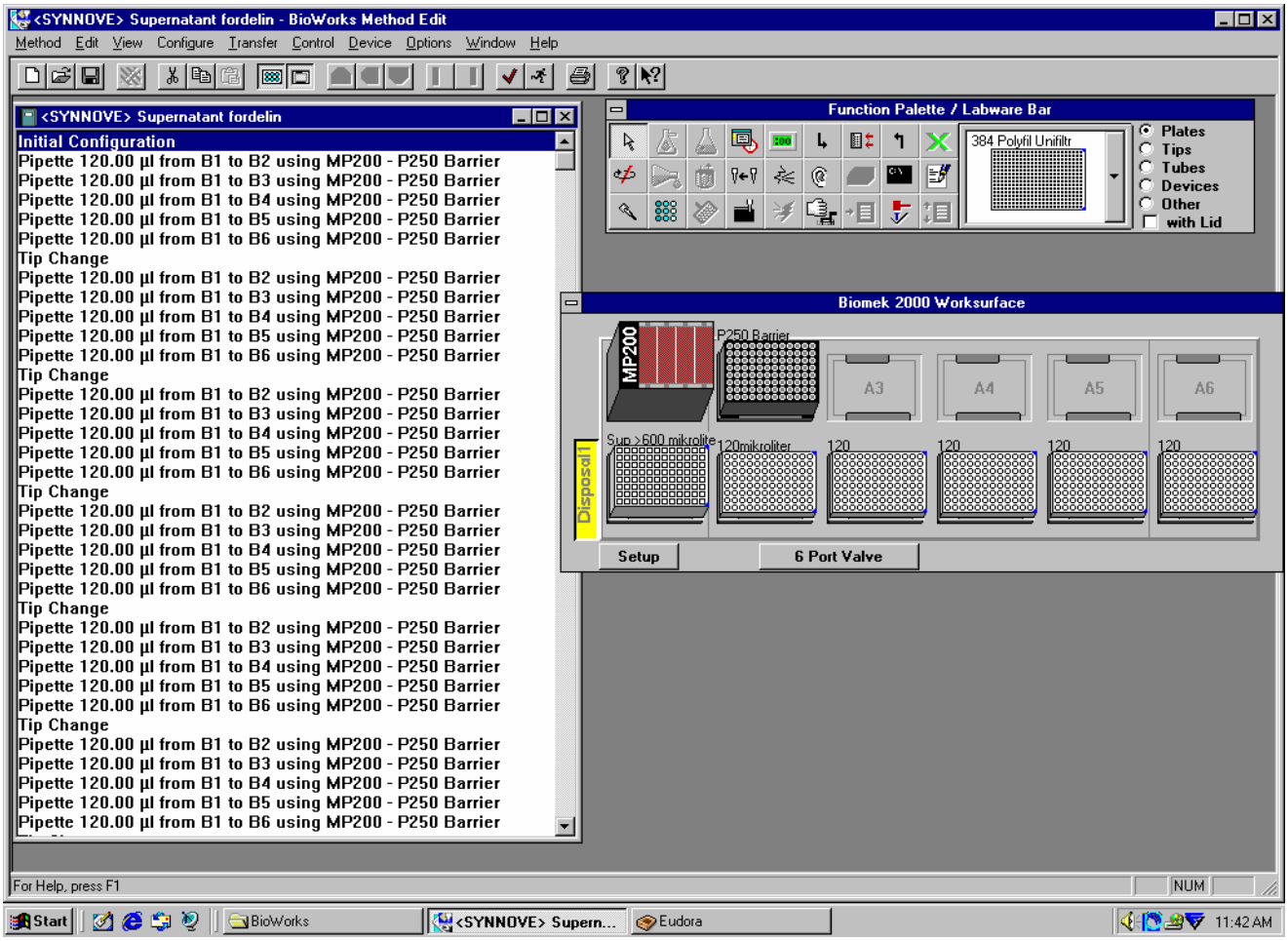
Script 1

BioWorks script giving Biomek 2000 instructions to pipete phage supernatant and a PCR mix to a PCR reaction plate. The figure also include the configuration of the Biomek 2000 workbench surface. This program was designed and used to set up the PCR reaction applied for production of sequencing templates of affinity selected phage single clones in a semi-automated fashion.



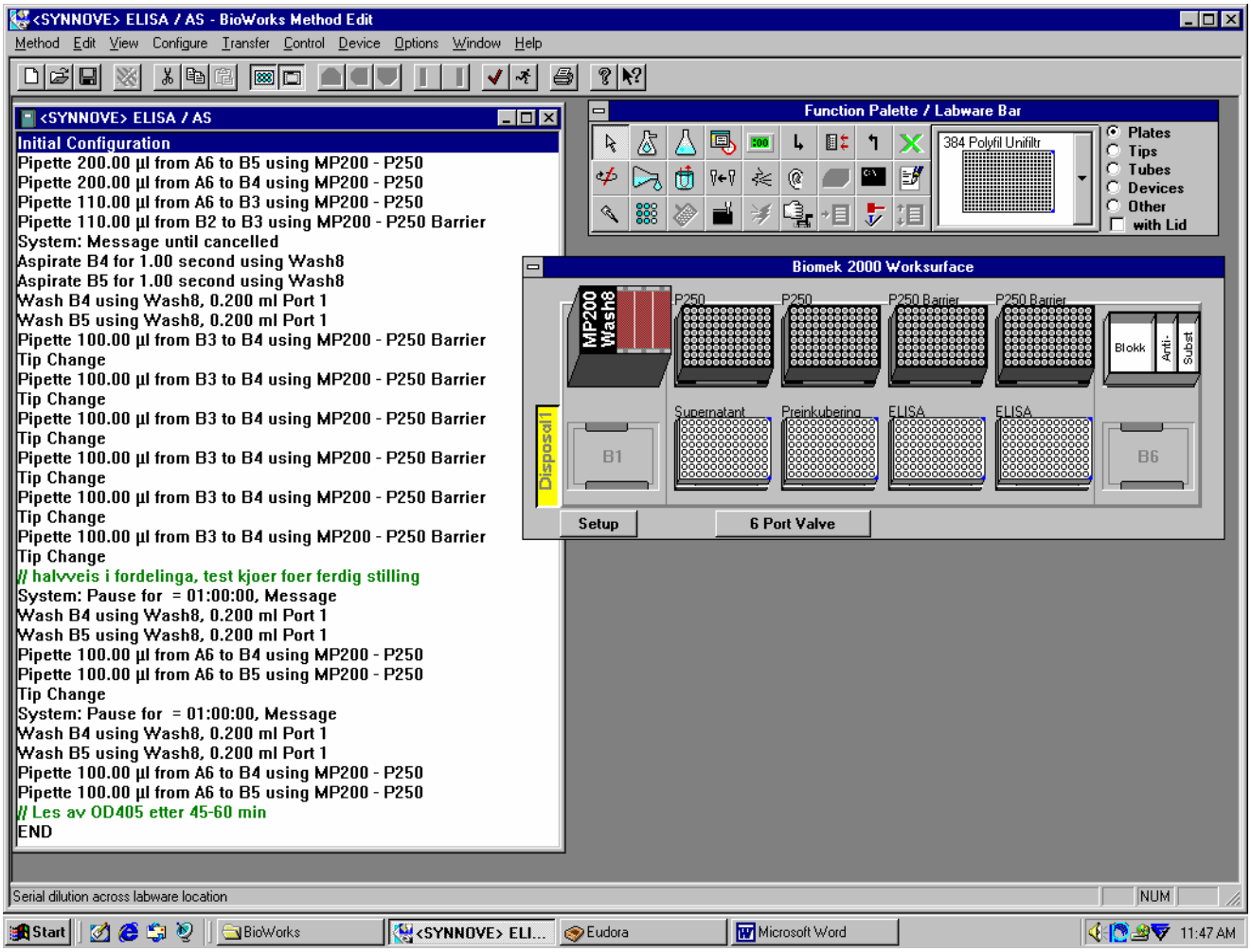
Script 2

BioWorks script giving Biomek 2000 instructions to pipete phage supernatant from a deep well culture growth plate and aliqoute it to 200µL V-bottom storage plates. The figure also include the configuration of the Biomek 2000 workbench surface. This program was designed to allow for semi automated handeling of supernatant from large numbers of selected phage single clones. The procedure was not implemented byond test runnings.



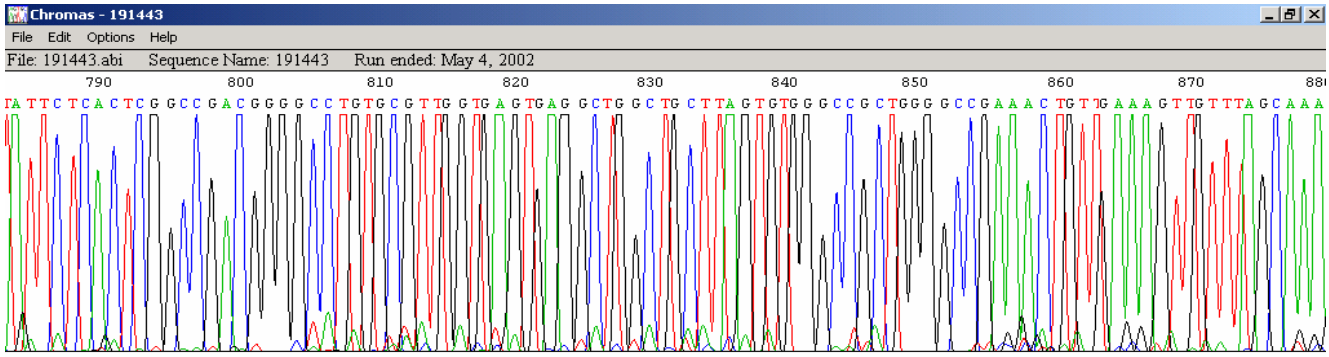
Script 3

BioWorks script giving Biomek 2000 instructions to set up a phage single clone ELISA. The figure also include the configuration of the Biomek 2000 workbench surface. This program was designed for robotic pipeting aid when screening phage single cones for target binding. ELISA plates were precoated. Only half of the actual script is shown her. The procedure was not implemented due to technical difficulties with the whashing tool.

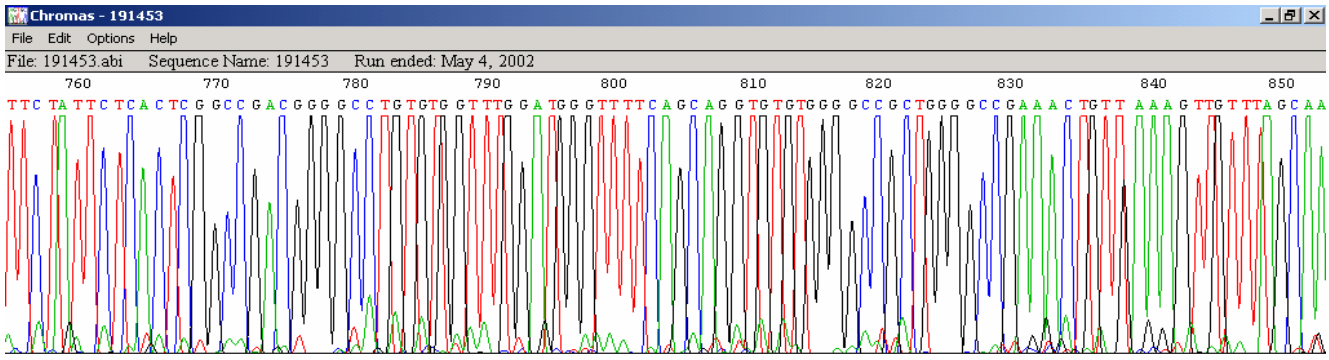


Appendix 2

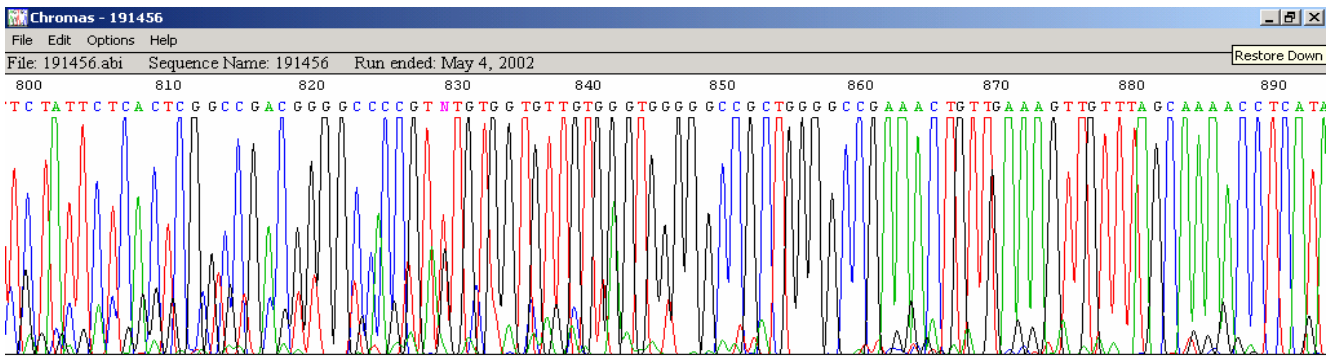
Abi files showing the sequence of the peptides displayed by phage recovered from the affinity selection procedure as presented in table 5. The sequence of interest is flanked by *Sfi*I restriction sites (GGGGCC).



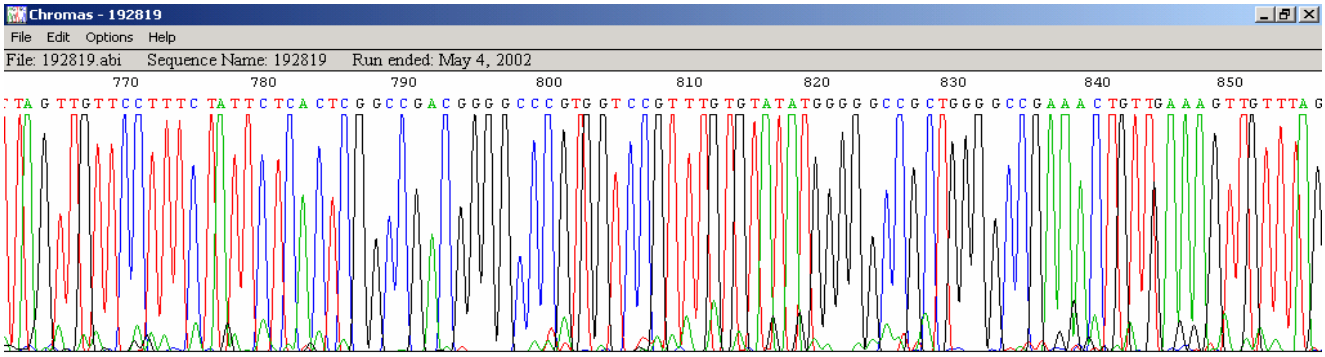
Amino acid sequence: CALVSEAGCLVWAA
Comment: The sequence has some unexpected features. G829 is represented by a somewhat low peak. If this is a "false" nucleotide, the sequence will harbor a TAG (Stop) codon prior to the expected final amino acid, TGT (Cystein). This appears to be unlikely. However the phage is clearly infectious, thus the peptide seems to be a 14-mer inframe without the cystein constriction. This sequence was given the name ASR.



Amino acid sequence: CVVWMGFQQVC
Comment: This sequence is identical to the previously identified clone named C9-4.

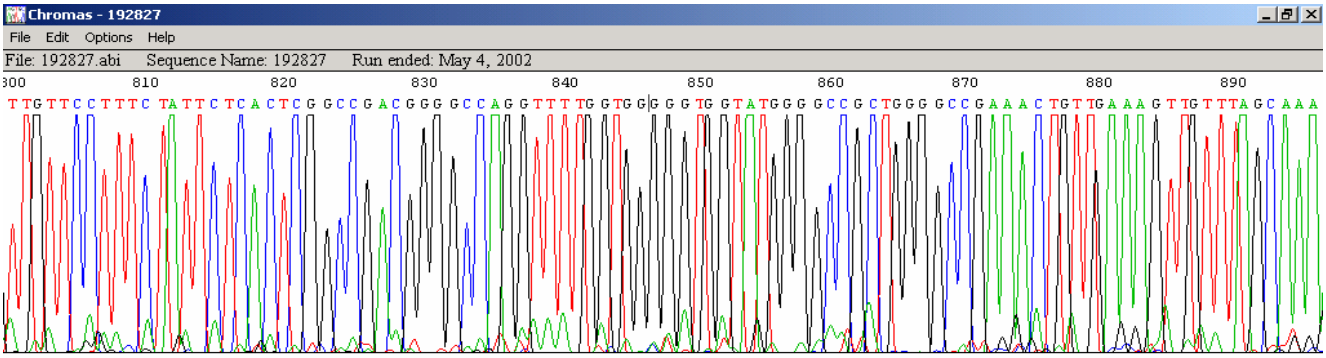


Amino acid sequence: PFVVLLV

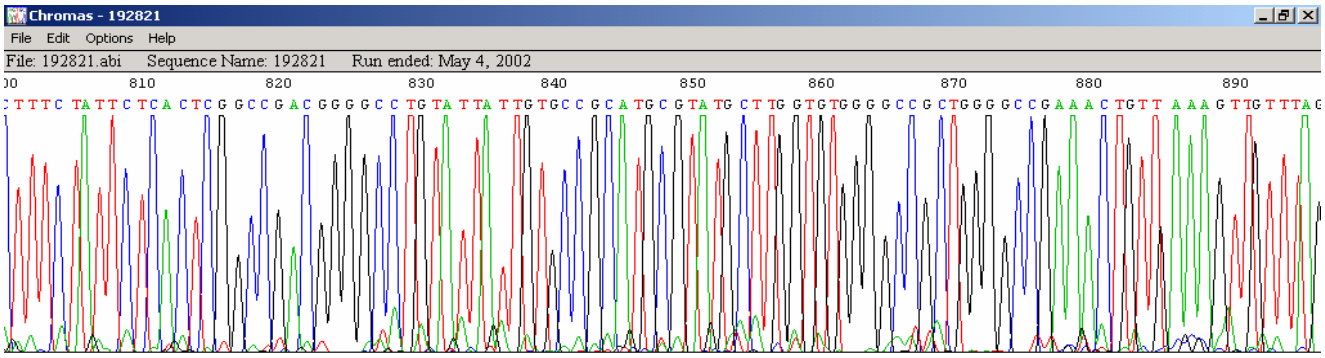


Comment: This is an unconstrained 7-mer peptide. Most likely a contaminant from the library screened in parallel to the Cys9 library.

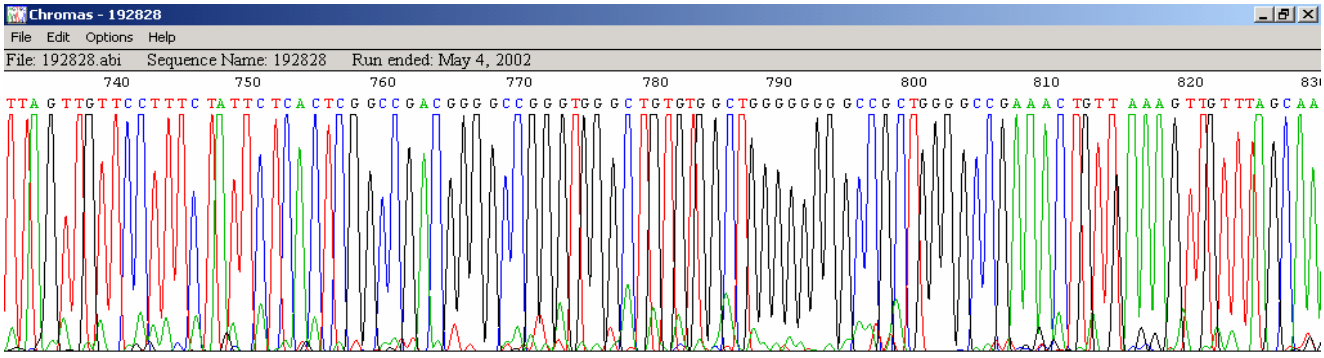
Amino acid sequence: RGPFVYM
Comment: This is an unconstrained 7-mer peptide. Most likely a contaminant from the library screened in parallel to the Cys9 library. This phage tested negative for SC binding in a confirmative ELISA.



Amino acid sequence: RFWWGWY
Comment: This is an unconstrained 7-mer peptide. Most likely a contaminant from the library screened in parallel to the Cys9 library.



Amino acid sequence: CIIVPHAYAWC



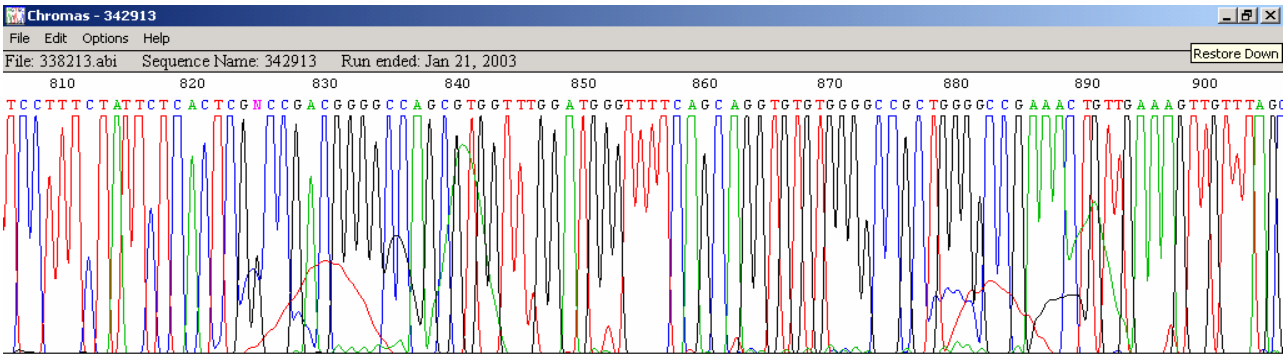
Amino acid sequence: GWAGWLG
Comment: This is an unconstrained 7-mer peptide. Most likely a contaminant from the library screened in parallel to the Cys9 library. This phage tested negative for SC binding in a confirmative ELISA.

Appendix 3

Abi files confirming the sequences of the peptide encoding inserts in the amino acid substitution mutation analysis. The sequence of interest is flanked by *Sfi*I restriction sites (GGGGCC). Serine substitutions are given in blue, whereas alanine substitutions are highlighted in red.

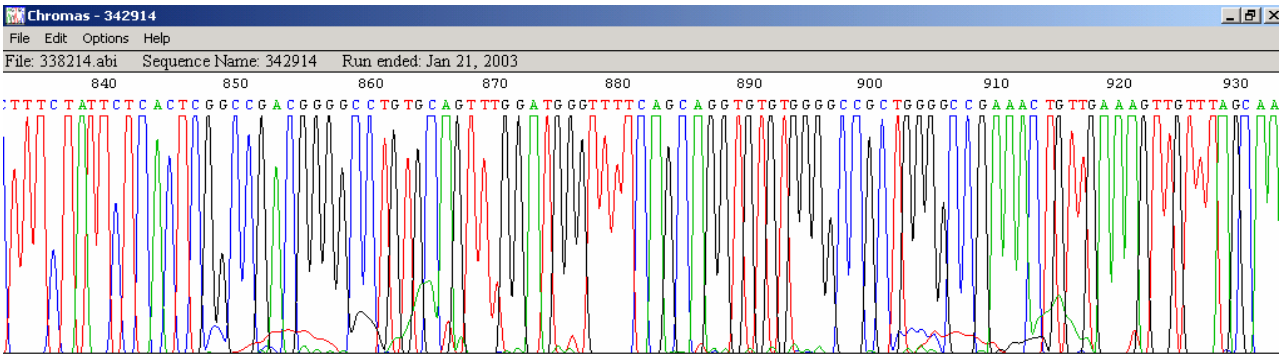
The prefix "AS" of the mutant's name conveniently means both "Amino acid Substitution" and "Anders Sandvik".

AS-C1S



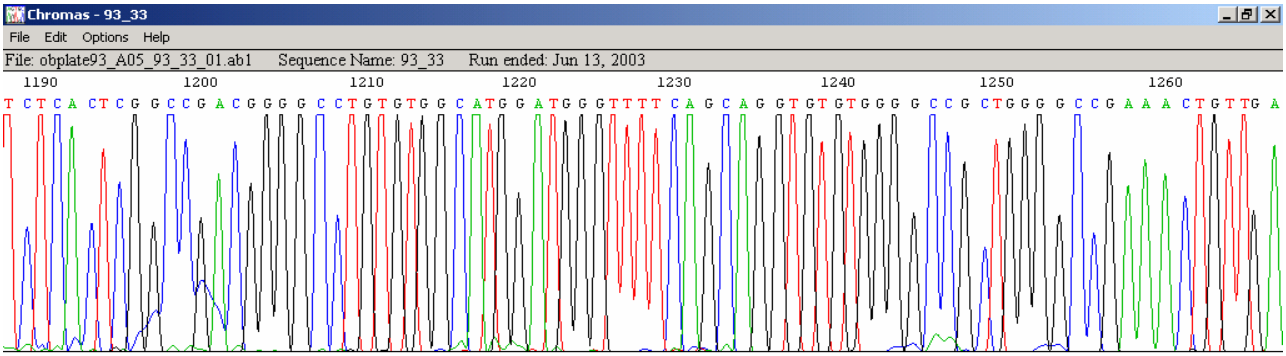
AGC GTG GTT TGG ATG GGT TTT CAG CAG GTG TGT

AS-V2A



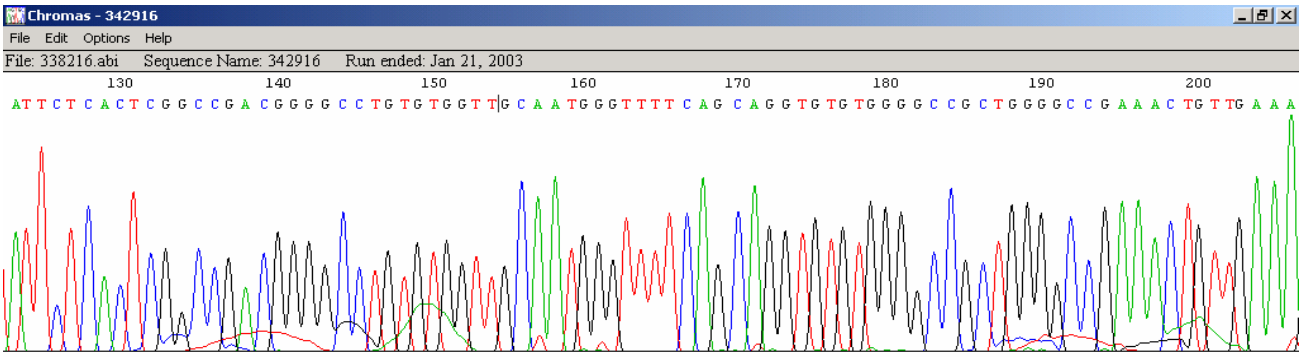
TGT GCA GTT TGG ATG GGT TTT CAG CAG GTG TGT

AS-V3A



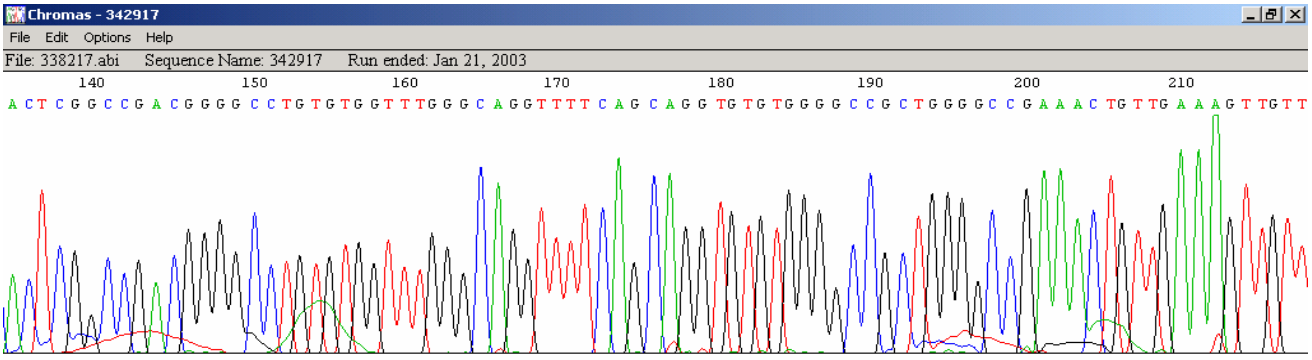
TGT GTG GCA TGG ATG GGT TTT CAG CAG GTG TGT

AS-W4A



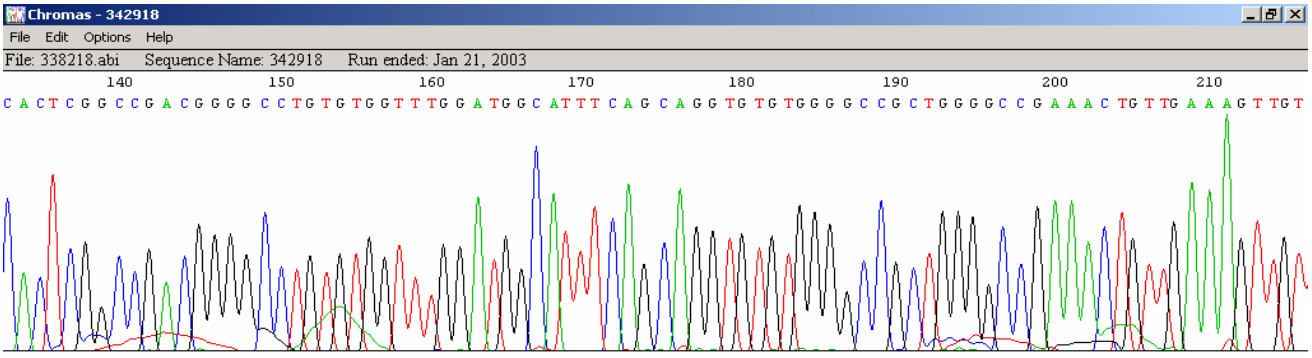
TGT GTG GTT GCA ATG GGT TTT CAG CAG GTG TGT

AS-M5A



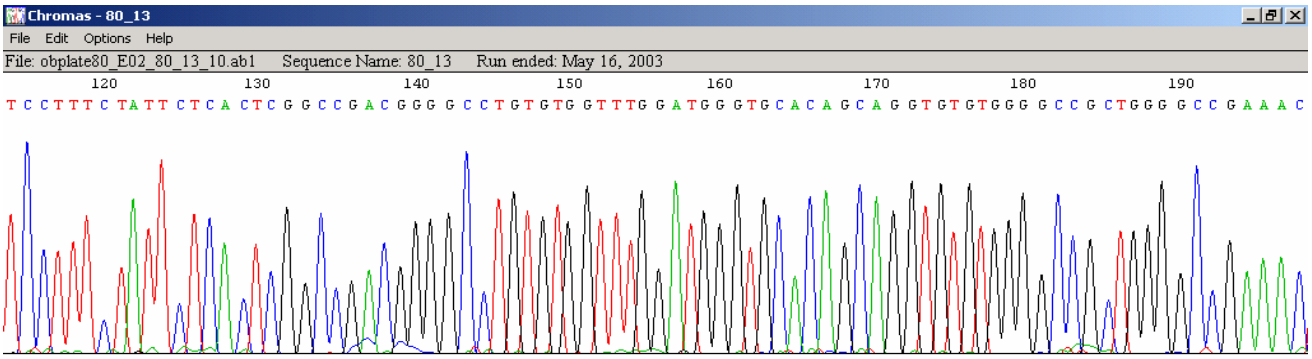
TGT GTG GTT TGG GCA GGT TTT CAG CAG GTG TGT

AS-G6A



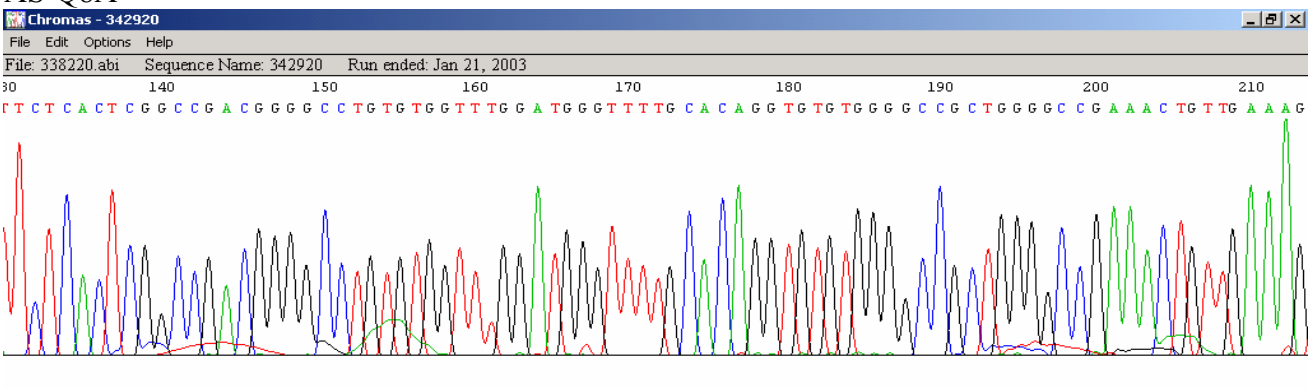
TGT GTG GTT TGG ATG GCA TTT CAG CAG GTG TGT

AS-F7A



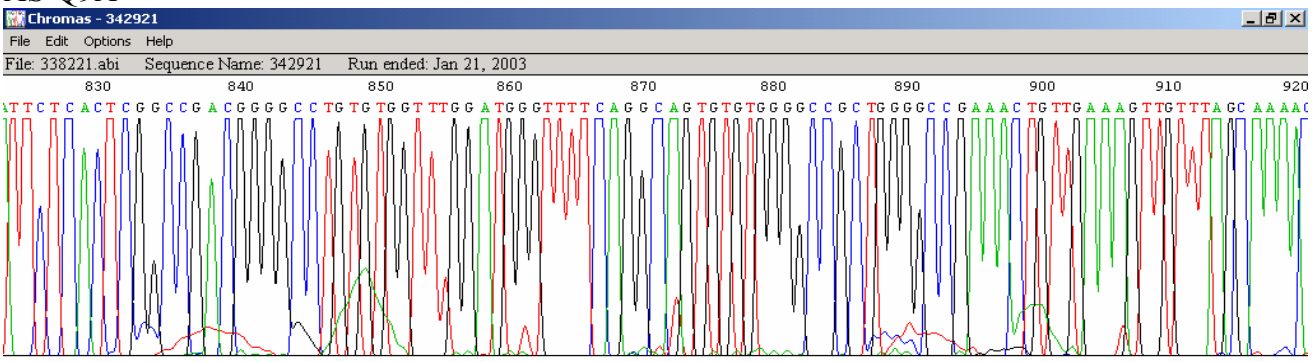
TGT GTG GTT TGG ATG GGT GCA CAG CAG GTG TGT

AS-Q8A



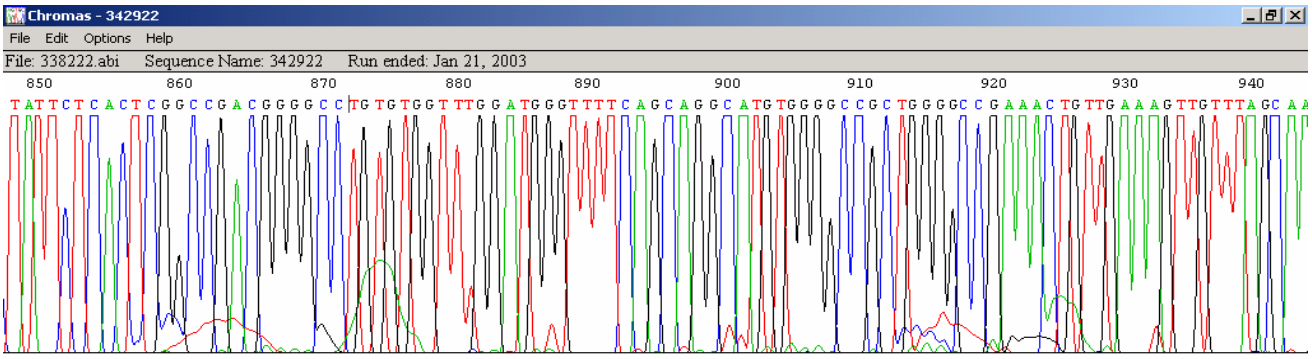
TGT GTG GTT TGG ATG GGT TTT GCA CAG GTG TGT

AS-Q9A



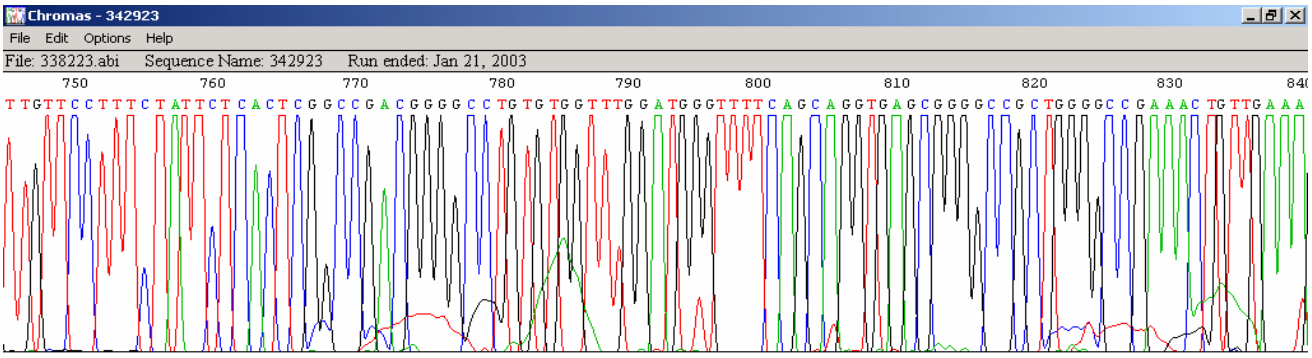
TGT GTG GTT TGG ATG GGT TTT CAG GCA GTG TGT

AS-V10A



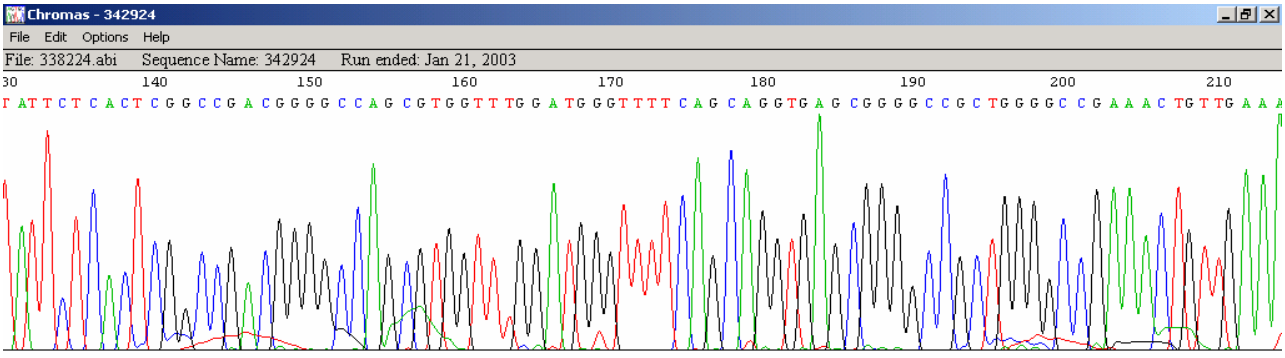
TGT GTG GTT TGG ATG GGT TTT CAG CAG GCA TGT

AS-C11S



TGT GTG GTT TGG ATG GGT TTT CAG CAG GTG AGC

AS-C1SC11S

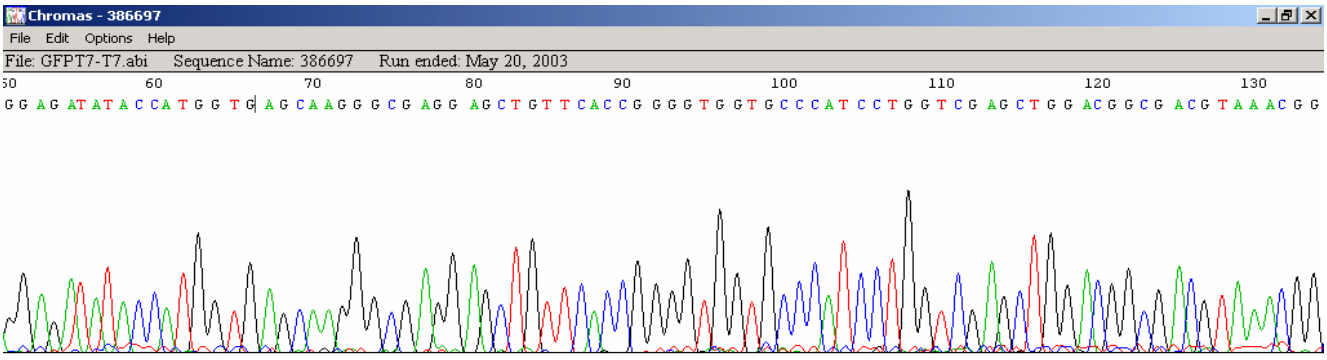


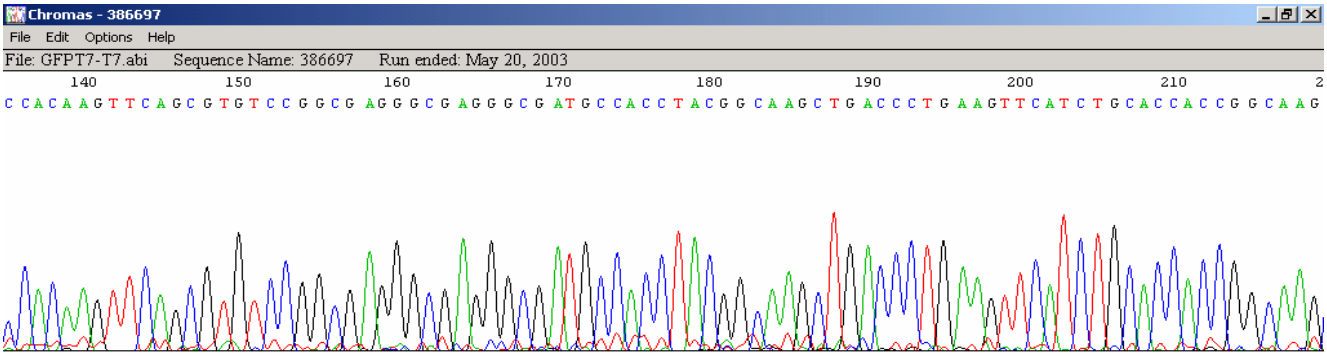
AGC GTG GTT TGG ATG GGT TTT CAG CAG GTG AGC

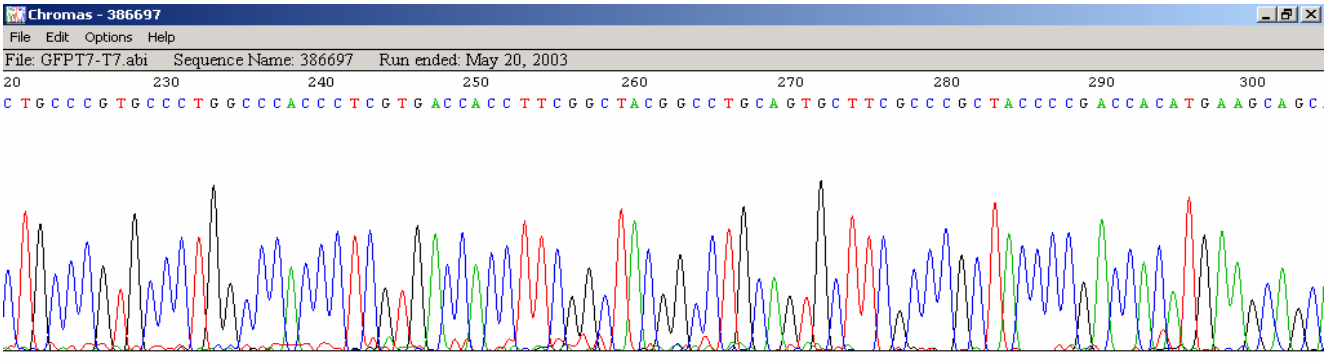
Appendix 4

Abi files presenting the detailed sequencing data from the GFP fusion work.
The sequences of interest are limited by *Nco*I (C˅CATGG) and *Xho*I (C˅TCGAG) restriction sites N-terminally and C-terminally, respectively.

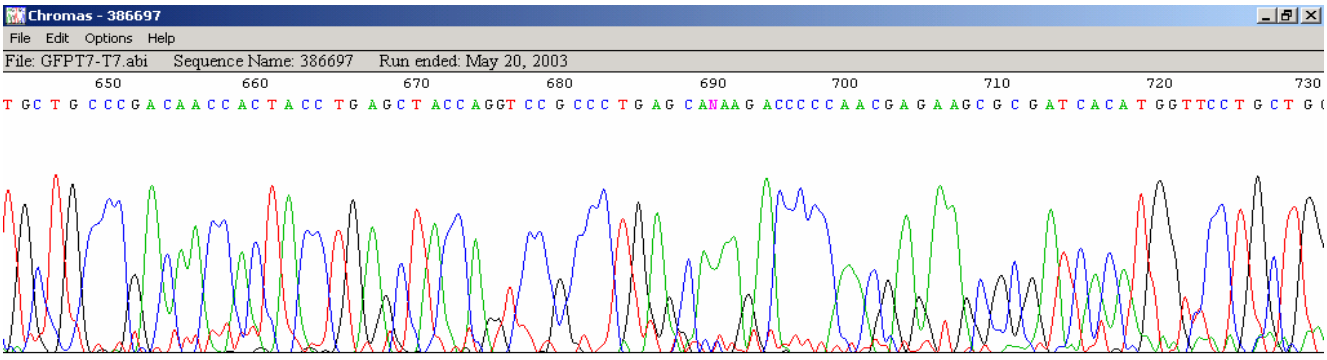
GFP without fusion peptide (GFP)
Comment: *Nco*I cuts at position 60 and the first codon of GFP (GTG) starts at position 64.
The GFP gene is terminated at position 787 with the *Xho*I restriction site.

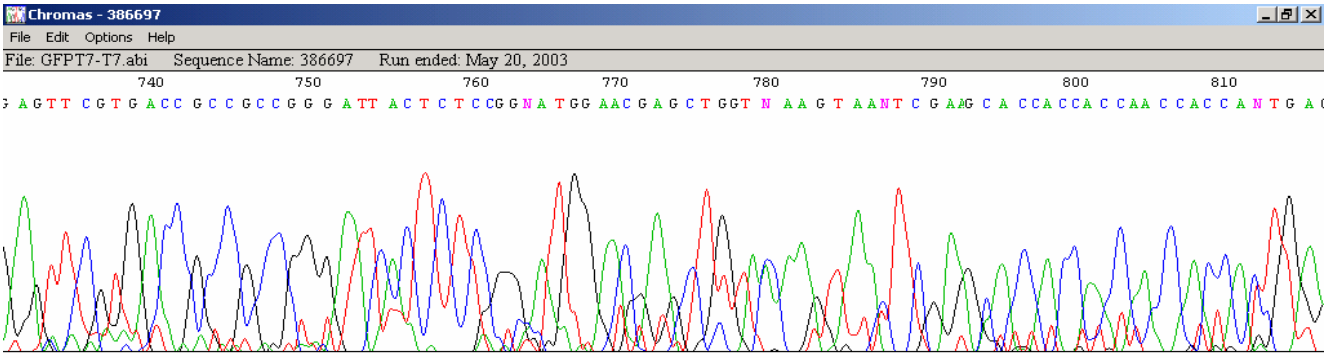






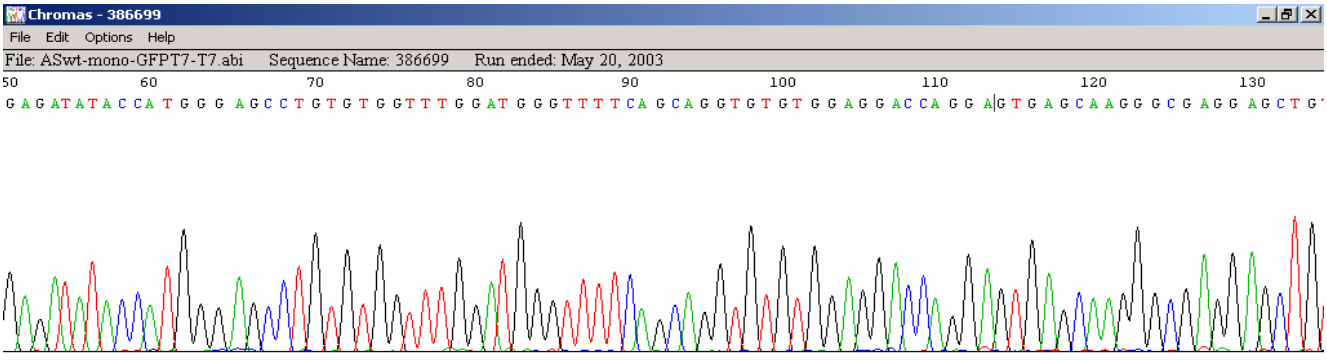

Chromas - 386697
File Edit Options Help
File: GFPT7-T7.abi Sequence Name: 386697 Run ended: May 20, 2003
310 320 330 340 350 360 370 380 390
A C G A C T T C T T C A A G T C C G C C A T G C C C G A A G G C T A C G T C C A G G A G C G C A C C A T C T T C T T C A A G G A C G A C G G C A A C T A C A A G A C C C G C
Chromas - 386697
File Edit Options Help
File: GFPT7-T7.abi Sequence Name: 386697 Run ended: May 20, 2003
400 410 420 430 440 450 460 470
G C C G A G G T G A A G T T C G A G G G C G A C A C C C T G G T G A A C C G C A T C G A G C T G A A G G G C A T C G A C T T C A A G G A G G A C G G C A A C A T C C T G G
Chromas - 386697
File Edit Options Help
File: GFPT7-T7.abi Sequence Name: 386697 Run ended: May 20, 2003
480 490 500 510 520 530 540 550
G G C A C A A G C T G G A G T A C A A C T A C A A C A G C C A C A A C G T C T A T A T C A T G G C C G A C A A G C A G A A G A A C G G C A T C A A G G T G A A C T T C A A
Chromas - 386697
File Edit Options Help
File: GFPT7-T7.abi Sequence Name: 386697 Run ended: May 20, 2003
560 570 580 590 600 610 620 630 640
G A T C C G C C A C A A C A T C G A G G A C G G C A G C G T G C A G C T C G C C G A C C A C T A C C A G C A G A A C A C C C C A T C G G C G A C G G T C C C G T G C

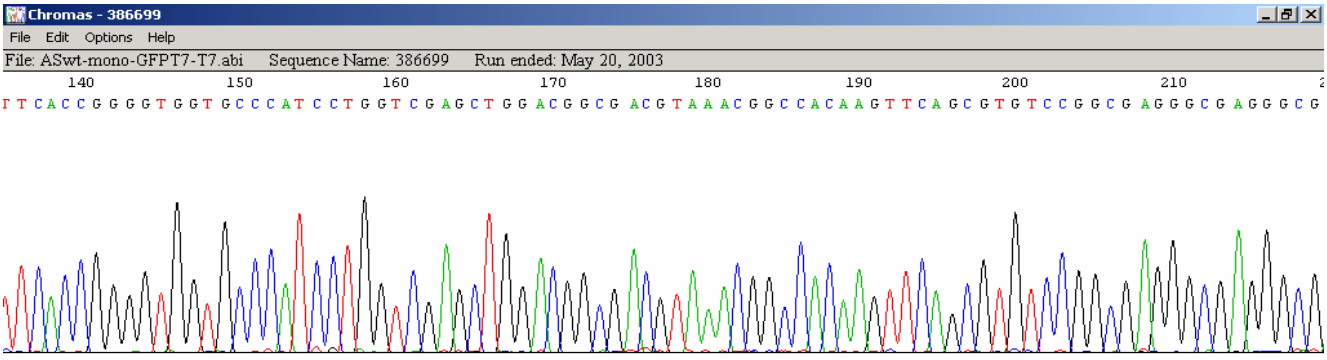


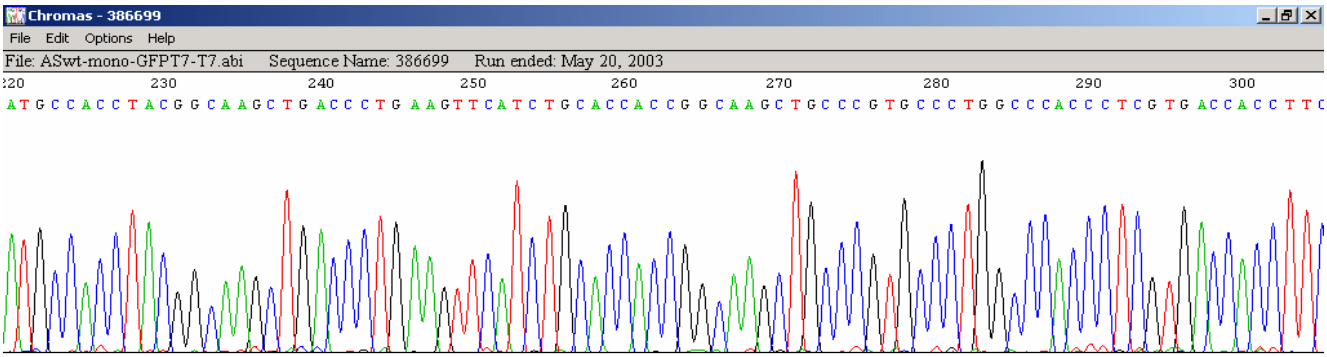


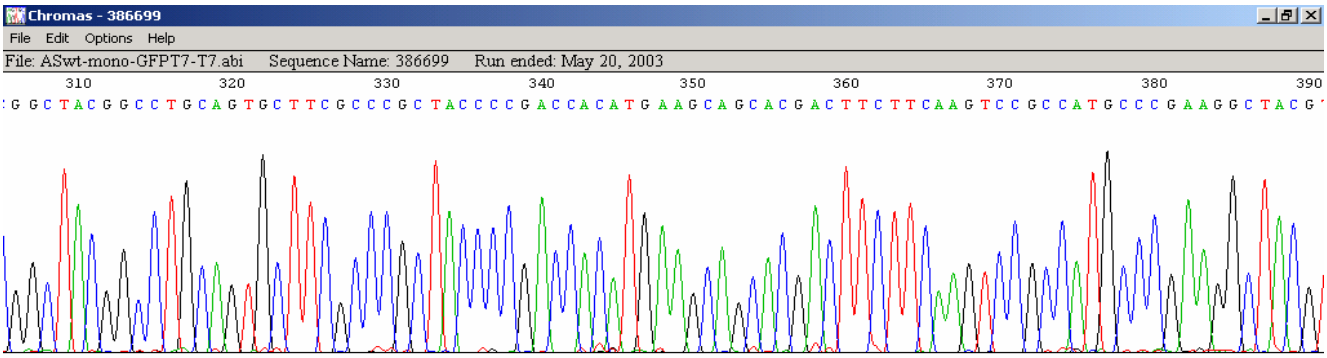
C9-4-Mono-GFP

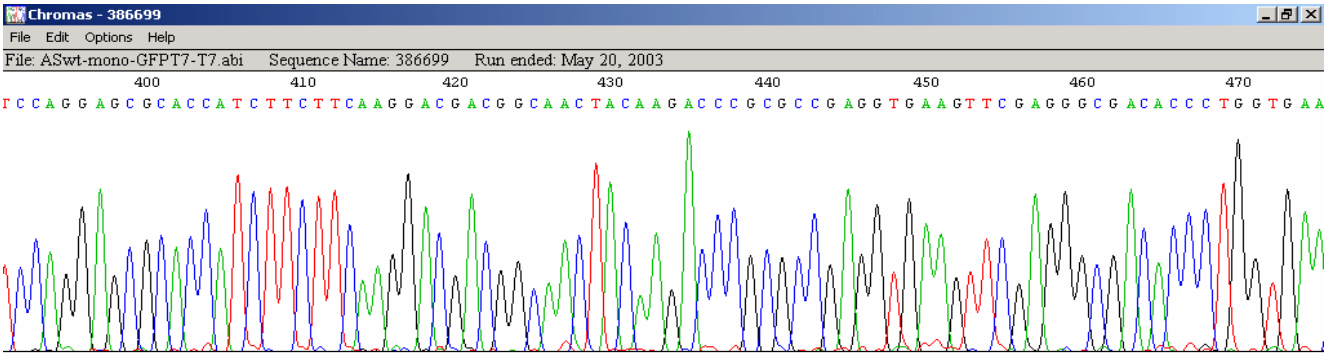
Comment: *Nco*I cuts at position 58 and the first codon of the SC binding peptide (TGT) starts at position 69. The GFP gene starts at (GTG) position 114 and is terminated at position 834with the *Xho*I restriction site.

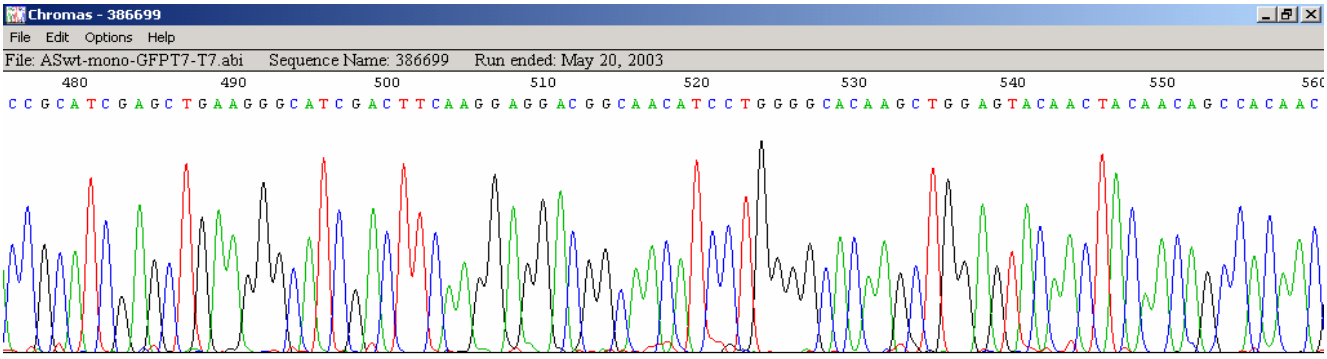


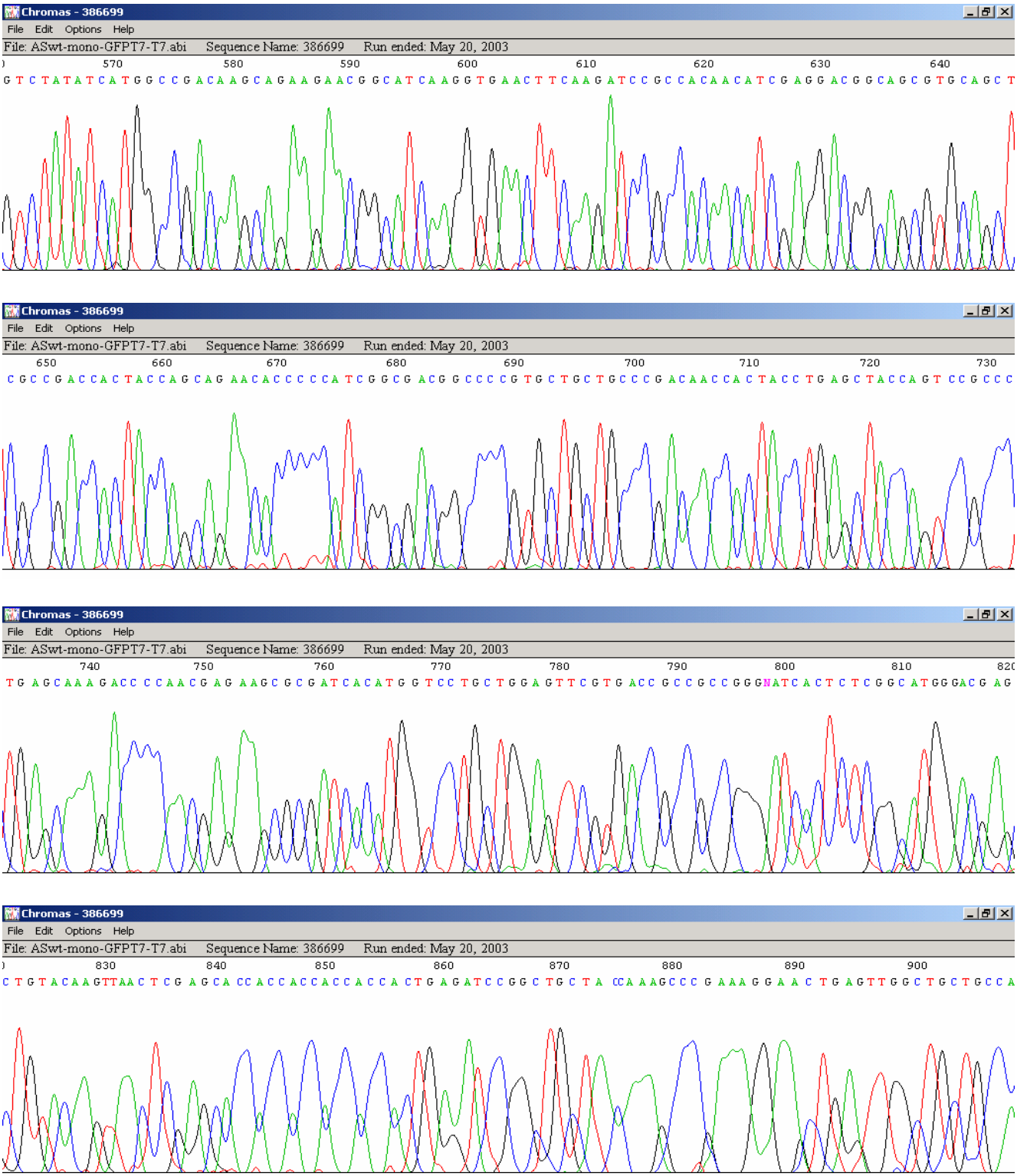



Chromas - 386699
File Edit Options Help
File: ASwt-mono-GFPT7-T7.abi Sequence Name: 386699 Run ended: May 20, 2003
220 230 240 250 260 270 280 290 300
ATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACCTTC


Chromas - 386699
File Edit Options Help
File: ASwt-mono-GFPT7-T7.abi Sequence Name: 386699 Run ended: May 20, 2003
310 320 330 340 350 360 370 380 390
GGCTACGGCCTGCAGTGCTTCGCCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCCCGAAGGCTACG


Chromas - 386699
File Edit Options Help
File: ASwt-mono-GFPT7-T7.abi Sequence Name: 386699 Run ended: May 20, 2003
400 410 420 430 440 450 460 470
TCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAA

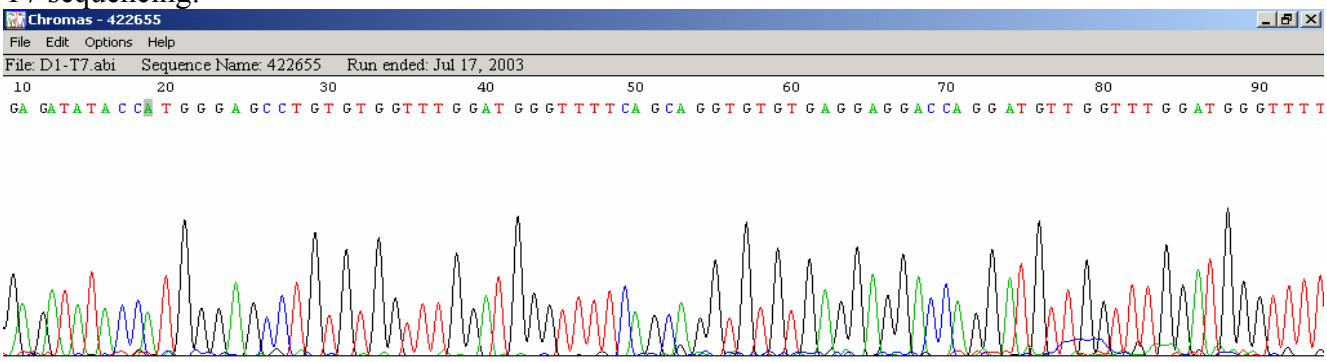

Chromas - 386699
File Edit Options Help
File: ASwt-mono-GFPT7-T7.abi Sequence Name: 386699 Run ended: May 20, 2003
480 490 500 510 520 530 540 550 560
CCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACTACAACAGCCACAAC

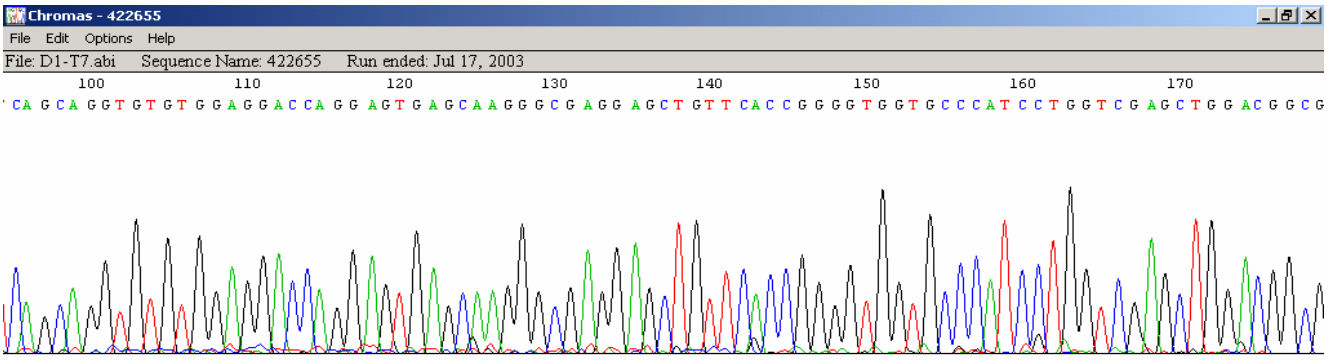

Chromas - 386699
File Edit Options Help
File: ASwt-mono-GFPT7-T7.abi Sequence Name: 386699 Run ended: May 20, 2003
570 580 590 600 610 620 630 640
G T C T A T A T C A T G G C C G A C A A G C A G A A G A A C G G C A T C A A G G T G A A C T T C A A G A T C C G C C A C A A C A T C G A G G A C G G C A G C G T G C A G C T
Chromas - 386699
File Edit Options Help
File: ASwt-mono-GFPT7-T7.abi Sequence Name: 386699 Run ended: May 20, 2003
650 660 670 680 690 700 710 720 730
C G C C G A C C A C T A C C A G C A G A A C A C C C C C A T C G G C G A C G G C C C C G T G C T G C T G C C C G A C A A C C A C T A C C T G A G C T A C C A G T C C G C C C
Chromas - 386699
File Edit Options Help
File: ASwt-mono-GFPT7-T7.abi Sequence Name: 386699 Run ended: May 20, 2003
740 750 760 770 780 790 800 810 820
T G A G C A A A G A C C C C A A C G A G A A G C G C G A T C A C A T G G T C C T G C T G G A G T T C G T G A C C G C C G C C G G G N A T C A C T C T C G G C A T G G G A C G A G
Chromas - 386699
File Edit Options Help
File: ASwt-mono-GFPT7-T7.abi Sequence Name: 386699 Run ended: May 20, 2003
830 840 850 860 870 880 890 900
C T G T A C A A G T T A A C T C G A G C A C C A C C A C C A C C A C C A C T G A G A T C C G G C T G C T A C C A A A G C C C G A A A G G A A C T G A G T T G G C T G C T G C C A

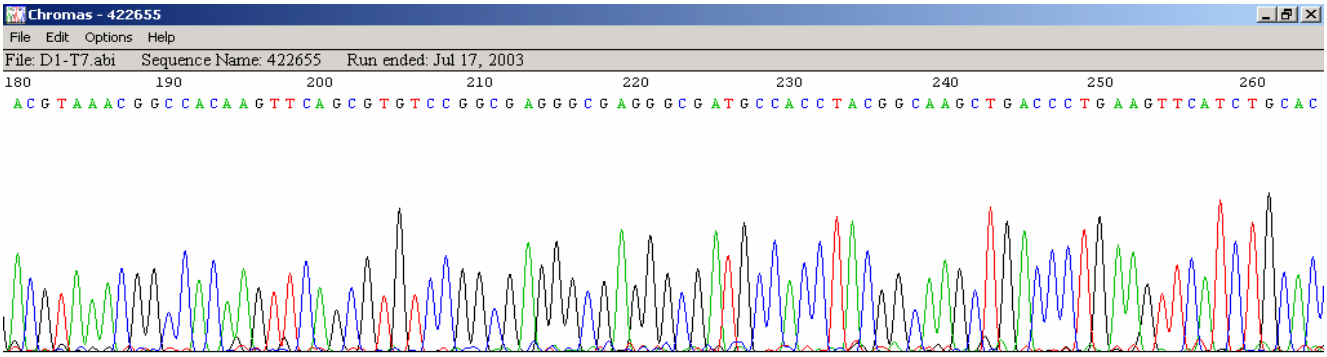
C9-4-Dimer-GFP

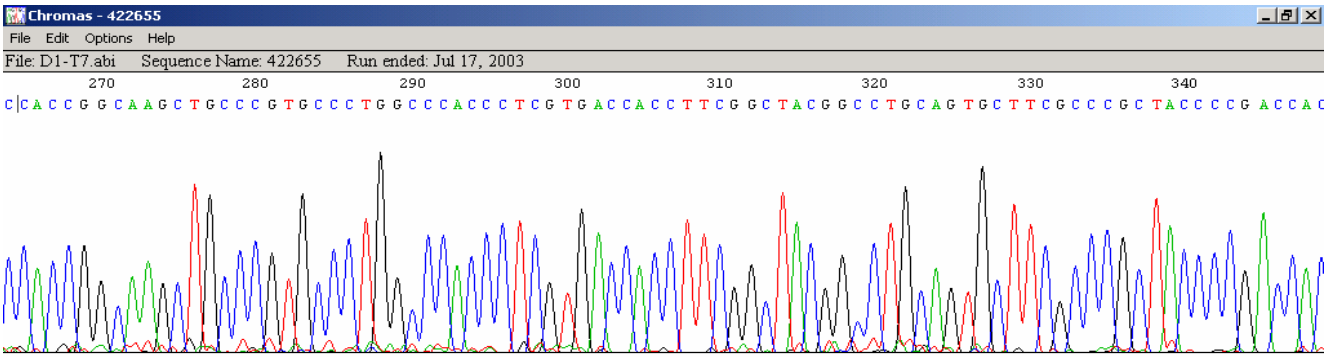
Comment: *Nco*I cuts at position 19 and the first codon of the first monomer of the SC binding peptide (TGT) starts at position 28. At positions 61 and 68, the sequence diverges from the expected one by lacking a G in each position. Consequently a frame shift mutation arises. Note that this insert was sequenced from both ends using primers T7 and pET-RP. The sequence obtained using pET-RP has been inverted to enable 5' to 3' reading for the complete sequence. The GFP gene starts at (GTG) position 119 on the T7 sequence. The GFP gene is terminated at position 876 on the pET-RP sequence with the *Xho*I restriction site. The sequences overlap to produce the complete C9-4-Dimer-GFP sequence. The thymine (T) at position 700 on the T7 sequence corresponds to the thymine in position 739 of the pET-RP sequence.

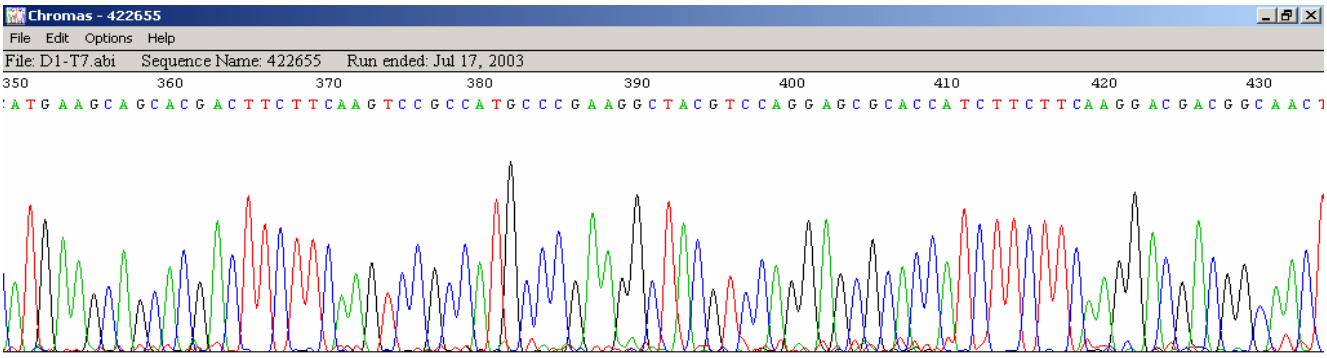
T7 sequencing:

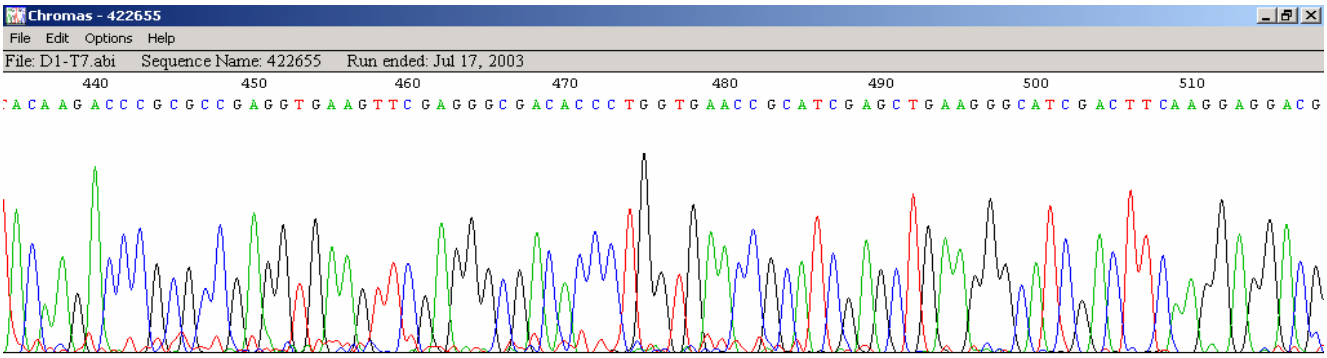


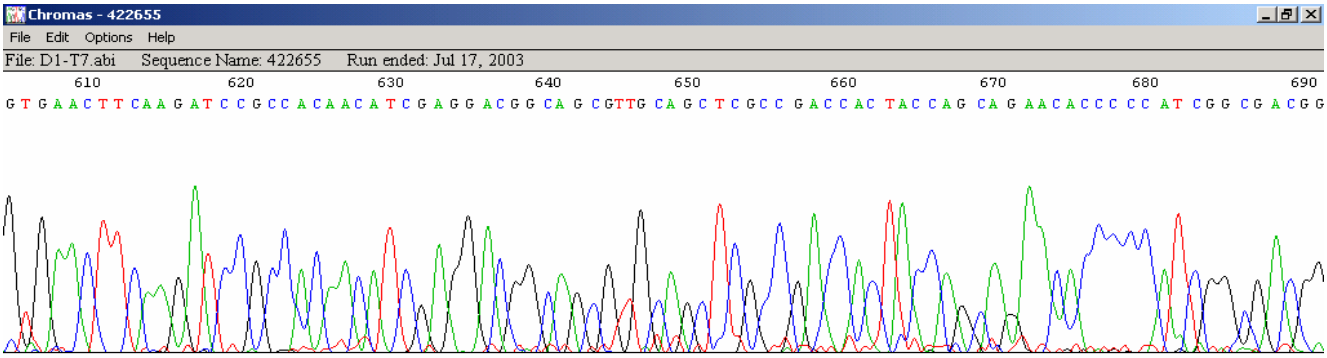


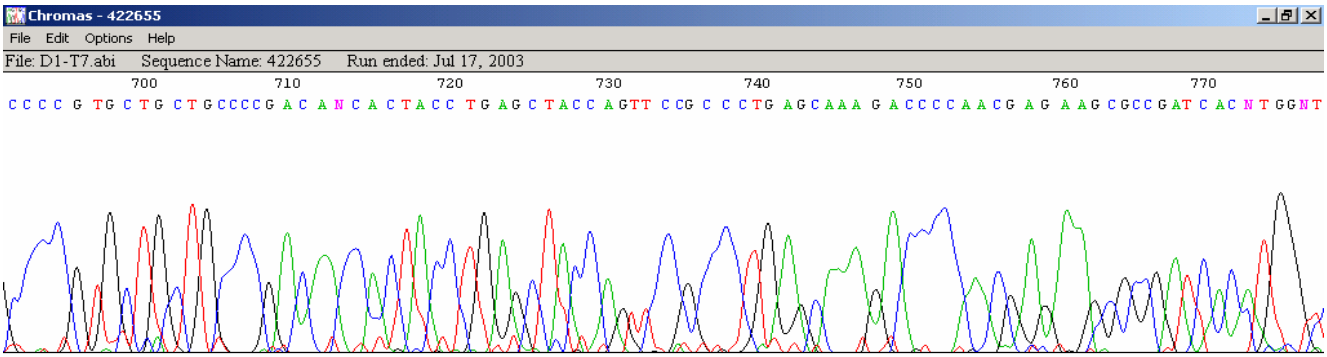



Chromas - 422655
File Edit Options Help
File: D1-T7.abi Sequence Name: 422655 Run ended: Jul 17, 2003
270 280 290 300 310 320 330 340

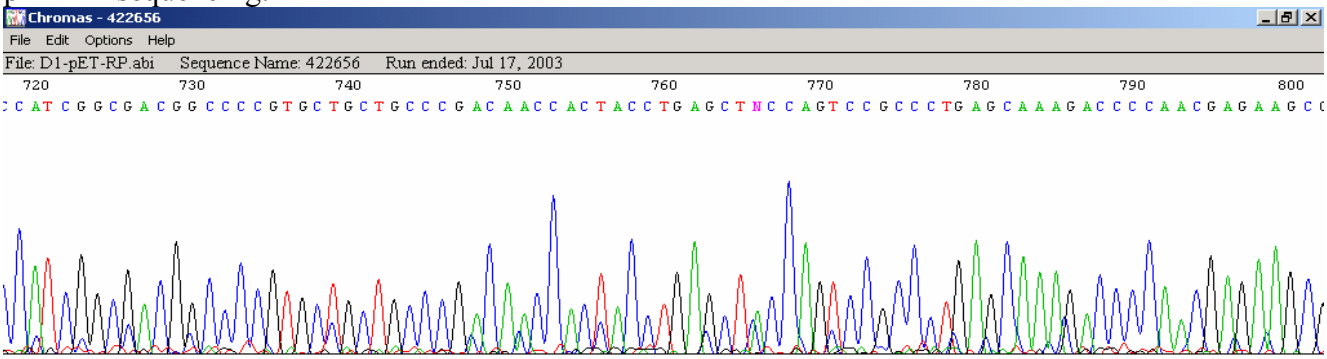

Chromas - 422655
File Edit Options Help
File: D1-T7.abi Sequence Name: 422655 Run ended: Jul 17, 2003
350 360 370 380 390 400 410 420 430

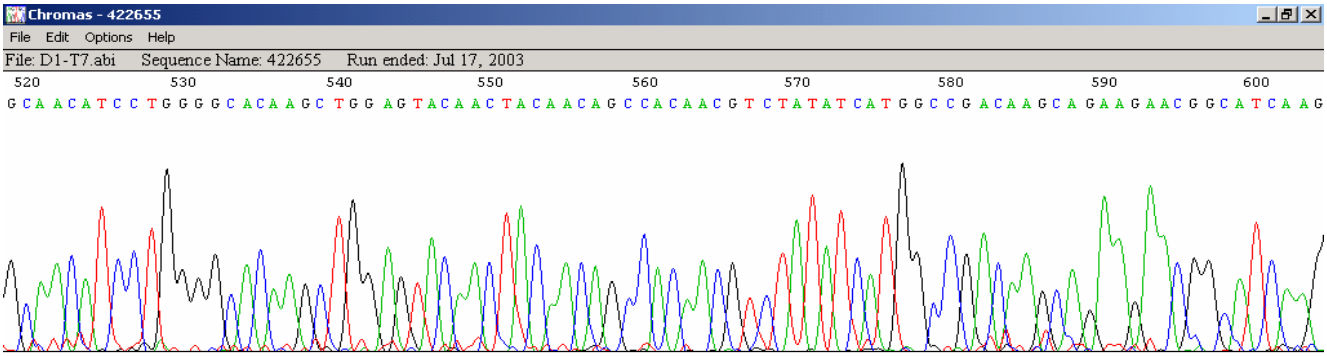

Chromas - 422655
File Edit Options Help
File: D1-T7.abi Sequence Name: 422655 Run ended: Jul 17, 2003
440 450 460 470 480 490 500 510

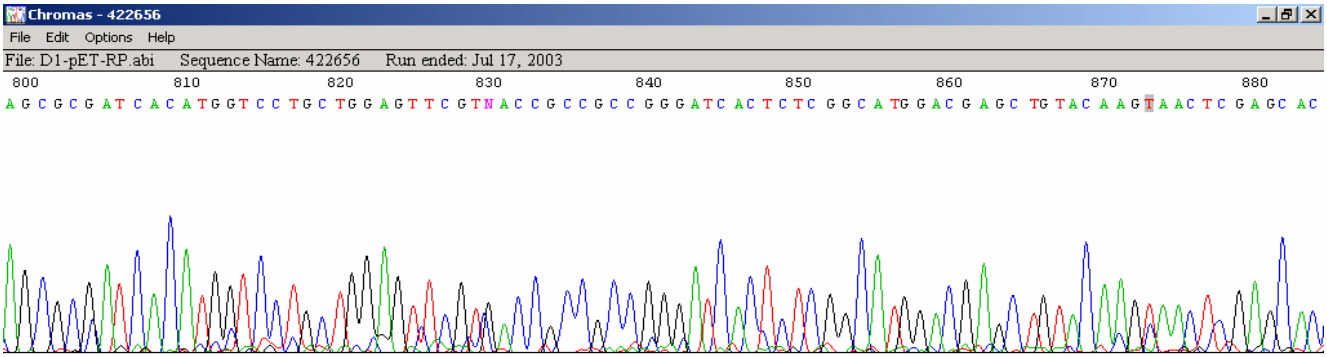

Chromas - 422655
File Edit Options Help
File: D1-T7.abi Sequence Name: 422655 Run ended: Jul 17, 2003
610 620 630 640 650 660 670 680 690



pET-RP sequencing:

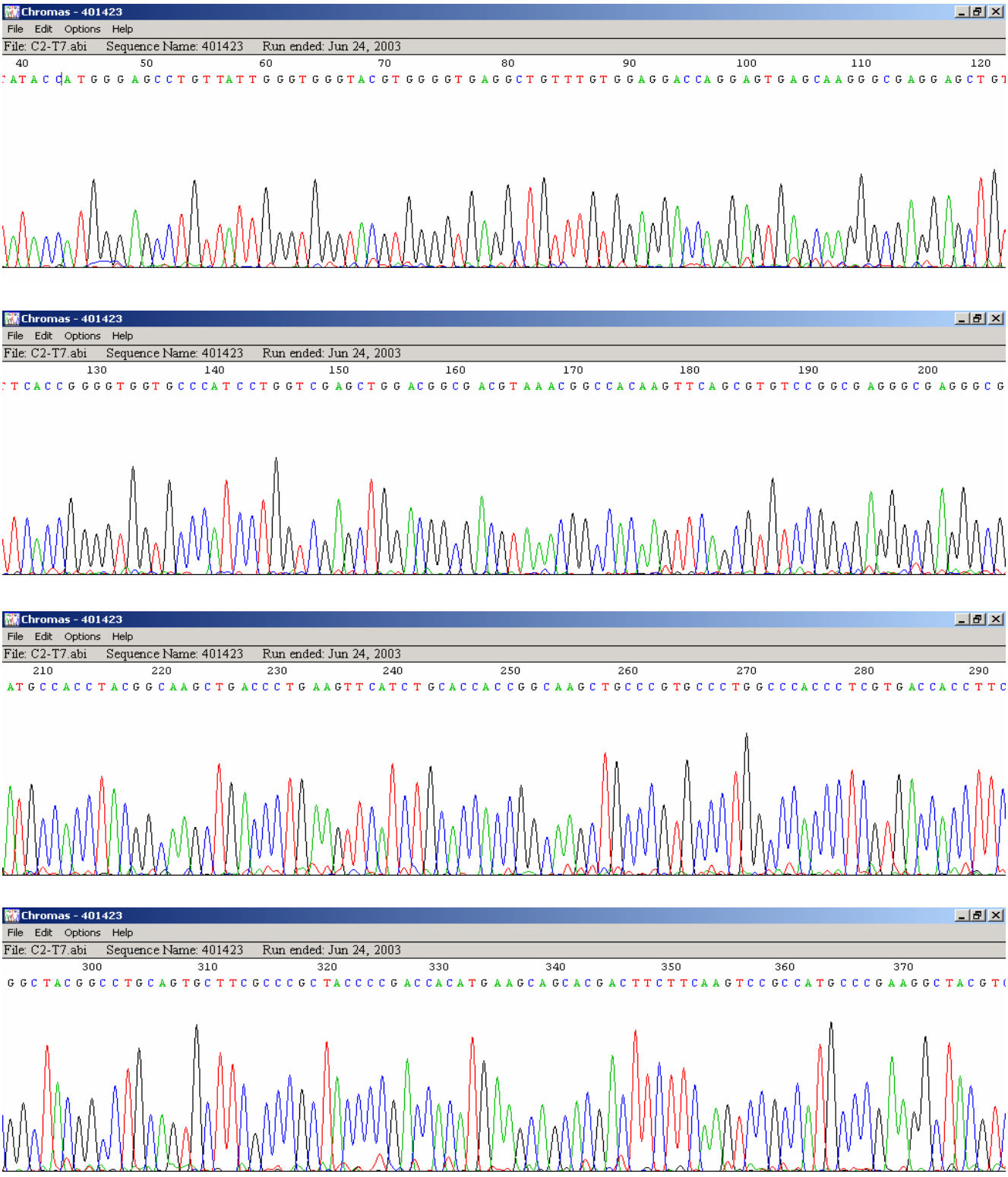




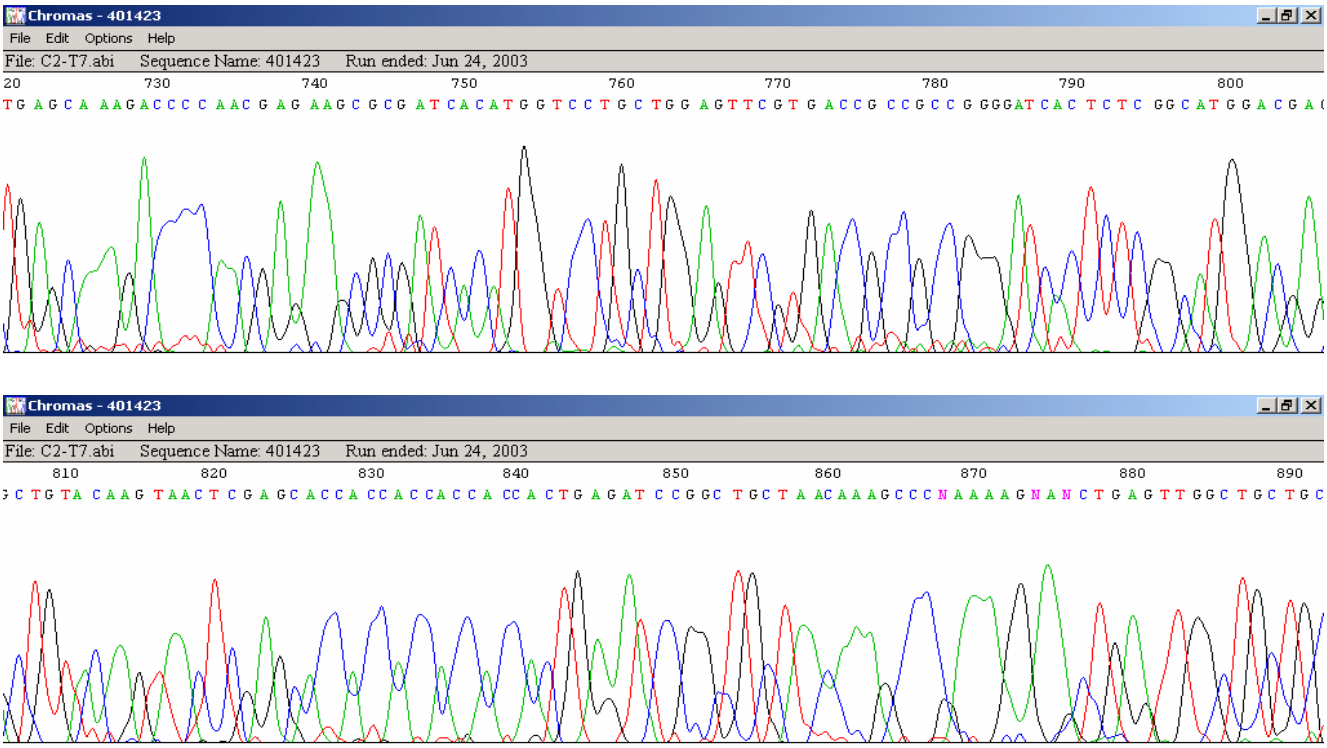


C1qC10-Mono-GFP

Comment: *Nco*I cuts at position 43 and the first codon of the C1q binding peptide (TGT) starts at position 53. The GFP gene starts at (GTG) position 101 and is terminated at position 820 with the *Xho*I restriction site.




Chromas - 401423
File Edit Options Help
File: C2-T7.abi Sequence Name: 401423 Run ended: Jun 24, 2003
380 390 400 410 420 430 440 450 460
C A G G A G C G C A C C A T C T T C T T C A A G G A C G A C G G C A A C T A C A A G A C C C G C G C C G A G G T G A A G T T C G A G G G C G A C A C C C T G G T G A A C C
Chromas - 401423
File Edit Options Help
File: C2-T7.abi Sequence Name: 401423 Run ended: Jun 24, 2003
470 480 490 500 510 520 530 540
G C A T C G A G C T G A A G G G C A T C G A C T T C A A G G A G G A C G G C A A C A T C C T G G G G C A C A A G C T G G A G T A C A A C T A C A A C A G C C A C A A C G T
Chromas - 401423
File Edit Options Help
File: C2-T7.abi Sequence Name: 401423 Run ended: Jun 24, 2003
550 560 570 580 590 600 610 620 630
C T A T A T C A T G G C C G A C A A G C A G A A G A A C G G C A T C A A G G T G A A C T T C A A G A T C C G C C A C A A C A T C G A G G A C G G C A G C G T G C A G C T C
Chromas - 401423
File Edit Options Help
File: C2-T7.abi Sequence Name: 401423 Run ended: Jun 24, 2003
640 650 660 670 680 690 700 710
G C C G A C C A C T A C C A G C A G A A C A C C C C A T C G G C G A C G G C C C C G T G C T G C T G C C C G A C A A C C A C T A C C T G A G C T A C C A G T C C G C C C


Chromas - 401423
File Edit Options Help
File: C2-T7.abi Sequence Name: 401423 Run ended: Jun 24, 2003
20 730 740 750 760 770 780 790 800
Chromas - 401423
File Edit Options Help
File: C2-T7.abi Sequence Name: 401423 Run ended: Jun 24, 2003
810 820 830 840 850 860 870 880 890